Yadav et al | Article 11693
J Pure Appl Microbiol. 2026. doi: 10.22207/JPAM.20.4.31
Received: 21 April 2026 | Accepted: 28 September 2026

P-ISSN: 0973-7510; E-ISSN: 2581-690X

RESEARCH ARTICLE OPEN ACCESS

Culture-Based Characterization and 16S rRNA Amplicon Profiling of Petroleum Hydrocarbon-Degrading Bacteria in Petroleum Sludge from Industrial Sites in Uttar Pradesh, India

Hitendra Yadav[1], Mohammed Haris Siddiqui[3] and Manish Dwivedi[1,2]*

[1]Amity Institute of Biotechnology, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, Uttar Pradesh, India.
[2]Research Cell, Amity University Uttar Pradesh, Lucknow Campus, Lucknow, Uttar Pradesh, India.
[3]Department of Bioengineering, Integral University, Lucknow, Uttar Pradesh, India.

Abstract

Petroleum hydrocarbons (PHCs) are hazardous pollutants released during extraction, refining, and transportation. This study combined culture-based enrichment with 16S rRNA gene amplicon sequencing to characterize the bacterial communities in petroleum sludge from four industrial sites in Uttar Pradesh, India. A pooled mixed-soil comparator (MIX-SOIL), prepared from equal portions of soil from the four sites, represented a less selectively enriched community. Samples were enriched in Bushnell-Haas broth with crude oil as the sole carbon source, and morphologically distinct colonies were purified. Community DNA was analyzed by Illumina MiSeq sequencing of the V3-V4 region, followed by quality filtering, 97% operational taxonomic unit (OTU) clustering, taxonomic assignment, and diversity analysis. *Pseudomonadota* was the dominant phylum, and *Pseudomonas* was the most abundant genus in the petroleum sludge samples; *Acinetobacter, Rhodococcus, Brevundimonas* and *Bacillus* were also detected. MIX-SOIL had the highest alpha diversity (Chao1 456; Shannon 3.08; Simpson 0.87; Fisher's alpha 61), whereas petroleum sludge samples had lower values (Chao1 286-341; Shannon 1.73-2.13; Simpson 0.56-0.77; Fisher's alpha 25-34). These results indicate selective enrichment of taxa with documented hydrocarbon degradation potential in petroleum sludge. Because functional profiles were inferred from 16S rRNA data, they require confirmation by targeted gene assays or shotgun metagenomics.

Keywords: Actinomycetal, Alkane Monooxygenase, Biodegradation, Petroleum Hydrocarbons, *Pseudomonadota*, 16S rRNA Amplicon Sequencing

***Correspondence:** mdwivedi@lko.amity.edu

Citation: Yadav H, Siddiqui MH, Dwivedi M. Culture-Based Characterization and 16S rRna Amplicon Profiling of Petroleum Hydrocarbon-Degrading Bacteria in Petroleum Sludge from Industrial Sites in Uttar Pradesh, India. *J Pure Appl Microbiol.* Published online 27 November 2026. doi: 10.22207/JPAM.20.4.31

© The Author(s) 2026. **Open Access**. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License which permits unrestricted use, sharing, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

INTRODUCTION

Petroleum hydrocarbons (PHCs) are major environmental pollutants released during extraction, refining, transport, and other petroleum-related processes. Managing these complex hydrophobic mixtures is difficult because their chemical properties govern environmental persistence and bioavailability. PHCs enter marine, freshwater, and terrestrial environments through acute spills and chronic releases and can accumulate in soils, sediments, and biological tissues when natural attenuation is insufficient.[1-4]

Given their bioaccumulation and biomagnification in the food chain, the toxicological risks of PHCs are substantial. Components, such as polycyclic aromatic hydrocarbons and BTEX (benzene, toluene, ethylbenzene, and xylenes), have been reported as xenobiotics, mutagens, carcinogens, endocrine disruptors, and inducers of oxidative stress, which may result in cellular and hepatic toxicity, as well as teratogenic effects.[5-9] The chemical stability of PHCs, associated with delocalized π-electron systems and hydrophobic interactions, further limits their environmental degradation.

Biodegradation, a microbial mediated process, is an ecologically compatible and cost-effective detoxification method that relies on indigenous microorganisms capable of metabolizing hydrocarbons as their sole carbon and energy sources.[10,11] The oxidation and cleavage reactions involved are mediated by enzymes such as monooxygenases, dioxygenases, and dehydrogenases.[12-14] Biodegradation efficiency depends on physicochemical and biological factors, including pH, temperature, nutrient availability, and microbial enzymatic capabilities.[15-18] Biosurfactants, surface-active biomolecules produced by some bacteria, facilitate microbial degradation by increasing hydrocarbon bioavailability and lowering the interfacial tension between hydrocarbons and water.[19,20]

Advances in molecular and metagenomic tools have broadened our understanding of hydrocarbon-degrading microbial communities.[21] Metagenomics examines genetic material recovered directly from environmental samples and can reveal microbial diversity, functional potential, and gene-environment associations.[22,23] In 16S rRNA amplicon studies, metabolic functions are inferred rather than measured directly; therefore, predicted catabolic markers such as alkB (alkane monooxygenase), catA (catechol 1,2-dioxygenase), and nahA (naphthalene dioxygenase) require experimental validation.[24]

Members of *Pseudomonadota*, Actinomycetota and Bacillota are common in hydrocarbon-affected environments and associated with alkane-oxidation and aromatic-ring-cleavage pathways.[25] These findings highlight the dynamic response of microbial communities to hydrocarbon stress and the value of metagenomic approaches for investigating bioremediation-related functional potential.[26-30]

This study aimed to identify and profile the microbial groups involved in petroleum hydrocarbon biodegradation using both culture-based and metagenomic analyses, addressing a knowledge gap concerning hydrocarbon-biodegrading microbes at industrial sites in Uttar Pradesh, India. These challenges can be addressed by identifying dominant bacterial genera, including *Pseudomonas, Acinetobacter,* and *Rhodococcus*, and by using genomic profiling (16S rRNA gene amplicon sequencing) alongside culture-based approaches to better understand the enzymatic processes of microbial petroleum hydrocarbon biodegradation. This approach aims to clarify the microbial and functional diversity of petroleum-biodegrading microbes and inform the development of effective bioremediation strategies.

This study is distinct from previous work because it combines culture-based approaches with metagenomic analysis to explore microbial communities and enzymes involved in hydrocarbon biodegradation, even in the presence of industrial contamination. This combination is important because it increases knowledge of microbial resistance mechanisms to environmental stressors such as industrial contaminations, and supports designing biodegradation strategies to remediate petroleum-contaminated sites.

MATERIALS AND METHODS

Study area

Petroleum- and crude-oil-contaminated soils were collected from four industrial sites in

Uttar Pradesh, India: Bajrang Petrochemicals Pvt. Ltd., Kanpur Dehat (BPC-KD), Bharat Petroleum Corporation Ltd., Kidwai Nagar, Kanpur (BPCL-KN), Bharat Petroleum Corporation Ltd., Kalpi Road, Kanpur (BPCL-KR) and Indian Oil Corporation Ltd., Mathura (IOCL-MT). A pooled mixed-soil comparator (MIX-SOIL) was prepared from equal portions of soil from the four sites before DNA extraction.

Sample collection and enrichment of hydrocarbon-degrading bacteria

Approximately 1 g of petroleum sludge or crude oil-contaminated soil from each site was aseptically transferred to a sterile container and transported to the laboratory. For enrichment, 1 g of each sample was inoculated into 250 mL of sterile Bushnell-Haas broth containing 1 g of crude oil as the sole carbon source in a 3-L Erlenmeyer flask. Cultures were incubated at 37 °C and 120 rpm for seven days; 1 mL of the enriched culture was then transferred to fresh medium under the same conditions, and this step was repeated three times. MIX-SOIL was a pooled comparator rather than an independent fifth collection site.

Isolation and purification of bacterial strains

Isolates from the enrichment cultures were spread-plated on Bushnell-Haas agar (BHA) containing crude oil as the sole carbon source and incubated at 37 °C for 48 hrs. Morphologically distinct colonies were selected, purified by successive subculturing, and maintained on nutrient agar slants at 4 °C for short-term storage and as 20% glycerol stocks at -20 °C for long-term storage.

DNA extraction

Genomic DNA was extracted from pure bacterial isolates used for the 16S rRNA sequencing and subsequent OTU, diversity, and functional analyses using an Xploregen DNA extraction kit (Biokart India Pvt. Ltd., Bangalore, India), according to the manufacturer's protocol. Cell pellets were lysed with buffer C1; proteins and RNA were removed with buffers C2 and C3; and DNA was bound to a silica membrane column with buffer C4, washed with buffer C5, and eluted in 100 µl of buffer C6 (10 mM Tris-HCl, pH 8.5). DNA concentration and purity were assessed by NanoDrop spectrophotometry (A260/A280 ratio, 1.8-2.0), and integrity was confirmed by agarose gel electrophoresis. Extracted DNA was stored at -20 °C.

PCR amplification of the 16S rRNA gene

The DNA extracted from selected pure bacterial isolates was used for amplification and sequencing of the 16S rRNA gene for isolate-level taxonomic identification. These sequences were not included in the subsequent OTU clustering, diversity, or functional analyses. For community-level microbial profiling, DNA was independently extracted from the respective composite microbial samples, and the resulting 16S rRNA amplicon sequences were used for OTU generation and downstream analysis. The bacterial 16S rRNA V3-V4 region was amplified with primers 341F (5'-CCTACGGGNGGCWGCAG-3') and 806R (5'-GACTACHVGGGTATCTAATCC-3'). Each 25 µL PCR contained 40 ng of template DNA, 10 pM of each primer and 12.5 µL high-fidelity Taq master mix. Cycling comprised 95 °C for 2 min; 25 cycles of 95 °C for 15 sec, 60 °C for 15 sec, and 72 °C for 2 min; and a final extension at 72 °C for 10 min. Amplicons were checked on a 1.5% agarose gel, purified with AMPure beads, and quantified with a Qubit dsDNA high-sensitivity assay.

Library preparation and sequencing

Purified 16S rRNA amplicons underwent a secondary PCR with Illumina barcoded adapters for library preparation, were purified with AMPure beads, quantified, and were sequenced on an Illumina MiSeq platform using 2 × 300 bp paired-end V3-V4 chemistry. The resulting BCL data were demultiplexed into FASTQ files, and quality was assessed with FastQC (v0.11.9) and MultiQC (v1.10.1); only high-quality reads were retained for analysis.

Bioinformatics and metagenomic data processing

Bioinformatics analysis used the Biokart 16S pipeline. Quality-filtered paired-end reads underwent chimera screening and were clustered into operational taxonomic units (OTUs) at 97% sequence identity. Representative sequences were assigned taxonomy against the NCBI 16S rRNA reference database. Because a single V3-V4 amplicon does not generally provide reliable

species-level resolution, interpretation was restricted to the genus level, and any lower-rank pipeline label was treated only as a closest-reference annotation.

Diversity and core microbiome analysis

The OTU table was filtered using a minimum count of 4 at 20% prevalence and a 10% interquartile-range low-variance filter. Samples were rarefied to the minimum library size and scaled by total-sum scaling. Chao1, Shannon, Simpson, and Fisher's alpha indices were calculated for each sample. Bray-Curtis dissimilarity was visualized by principal coordinate analysis, and hierarchical clustering used Ward linkage. Because each site was represented by one composite sample, diversity values were interpreted descriptively, and no inferential test of site differences was performed. The core microbiome threshold was set at ≥25% prevalence and ≥1% relative abundance.

Table 1. Sequencing Depth and GC Content of Samples

Sample ID	%GC	Total Sequences (Million Reads)
BPC-KD	53.5	0.146
BPCL-KN	54.0	0.148
BPCL-KR	55.0	0.143
IOCL-MT	54.5	0.142
MIX-SOIL	53.0	0.147

Functional prediction analysis

Potential functional composition of the communities was inferred from representative 16S rRNA sequences using PICRUSt2 and summarized by KEGG Level-2 pathways. Predicted functions for alkane monooxygenase (alkB), naphthalene dioxygenase (nahA) and catechol 1,2-dioxygenase (catA) were examined. These outputs are computational predictions, and not direct measurements of genes, transcripts, or enzyme activities.

Statistical analysis

Sequencing depth, GC content, relative abundance, and alpha diversity indices were summarized descriptively. Bray-Curtis dissimilarities supported ordination and hierarchical clustering. Because each location was represented by a single composite sample (n = 1), one-way ANOVA and site-level significance testing were not appropriate and were not performed.

RESULTS

Sample sequencing metrics

All five libraries had similar sequencing yields (0.142-0.148 million reads) and GC contents

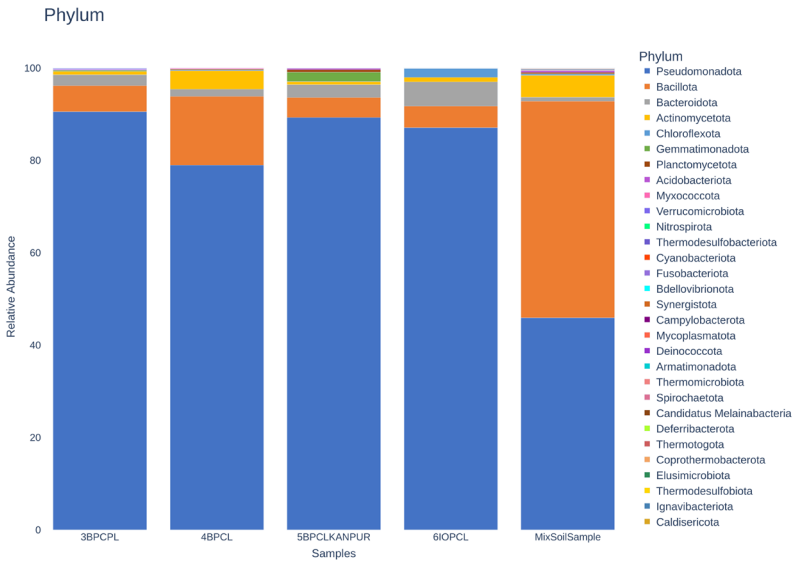

Figure 1. Phylum Level-Taxonomy Plot Analysis: representation of relative distribution of taxonomy of microorganisms in four petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT), as well as a mixed soil sample

(53.0%-55.0%), indicating comparable sequencing efforts (Table 1). BPCL-KN had the highest yield (0.148 million reads), and IOCL-MT had the lowest yield (0.142 million reads); MIX-SOIL had 53.0% GC content and 0.147 million reads.

Taxonomic analysis

Figure 1 shows the relative taxonomic distribution across the four petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT) and MIX-SOIL. In all four sludge samples, the highest proportion (80%-90%) represented *Pseudomonadota*, consistent with enrichment of taxa associated with hydrocarbon degradation.

In contrast, MIX-SOIL showed greater taxonomic diversity, with *Pseudomonadota* comprising 45%-55% of the community alongside higher proportions of Bacillota and other major phyla, indicating greater complexity. MIX-SOIL also had a higher abundance of low-abundance phyla, specifically Chloroflexota, Gemmatimonadota, and Planctomycetota, suggesting greater stability of the microbiota.

Although these phyla were present at low but stable frequencies in the petroleum sludge samples, their higher abundance in MIX-SOIL reflects a relatively diverse microbial community. Overall, the sludge samples mostly comprised hydrocarbon-decomposing microbes, whereas MIX-SOIL had a relatively balanced microbial community.

The community profiles of the four petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT) and one mixed soil sample (MIX-SOIL) at the class level are shown in Figure 2. The sludge samples were dominated by Gammaproteobacteria, ranging from 70%-90%, In MIX-SOIL, Gammaproteobacteria were lower, at approximately 40%-45%, with other classes such as Bacilli, Actinobacteria, and Clostridia predominant. This showed that Gammaproteobacteria, which are hydrocarbon degraders, were largely present in the petroleum sludge samples. The proportion of Gammaproteobacteria was lower in BPCL-KN, with higher proportions of Bacilli, Alphaproteobacteria, and Actinobacteria, indicating variation among the sludge samples. In MIX-SOIL, Gammaproteobacteria were lower, at approximately 40%-45%, whereas Bacilli, Actinobacteria, and Clostridia were more abundant. Minor classes such as Sphingobacteria, Planctomycetia, and Thermoleophilia were absent from the sludge but abundant in the mixed soil, reflecting a more complex microbial community.

Figure 3 shows the order-level microbial profiles across four petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT) and a mixed

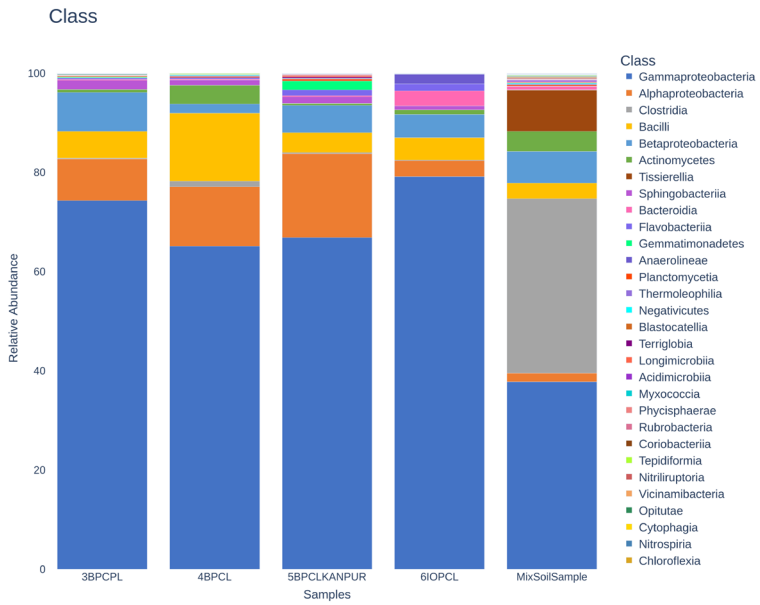

Figure 2. Class Level-Taxonomy Plot Analysis: The community profiles of four petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT) and one mixed soil sample (MIX-SOIL) at the class level

soil sample (MIX-SOIL). The petroleum sludge samples were dominated by Pseudomonadales, making up 60%-90% of the total abundance, consistent with a high relative abundance of the hydrocarbon-degrading genus *Pseudomonas*. Minor site-specific variations were evident in BPCL-KN and BPCL-KR, with a slight reduction in Pseudomonadales, allowing for an increased representation of *Bacillales*, *Burkholderiales*, *Rhizobiales*, and *Xanthomonadales*. In the mixed soil sample, Pseudomonadales was only ~15-20%, whereas a relatively high representation of soil-associated orders such as *Bacillales*, *Lactobacillales*, *Micrococcales*, and *Rhizobiales* reflected a more compositionally diverse community relative to the sludge samples. Overall, the petroleum sludge samples were strongly enriched in Pseudomonadales, whereas the mixed-soil samples presented a more balanced and diverse order of composition.

Figure 4 shows the families present in the petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT) along with a mixed sample of soil (MIX-SOIL). In the sludge

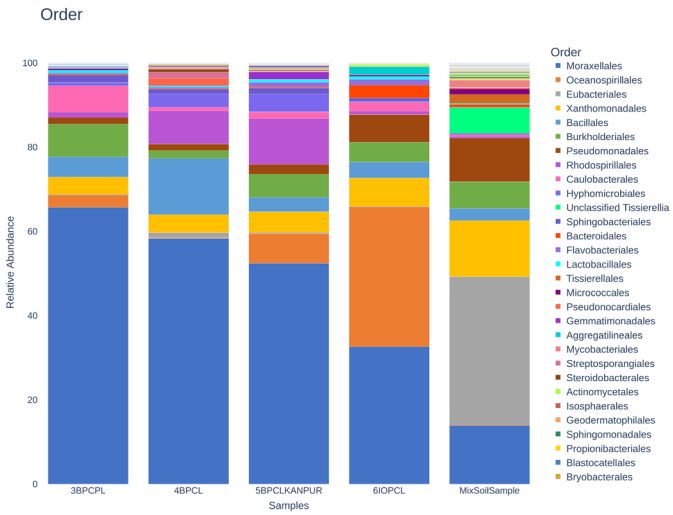

Figure 3. Order Level-Taxonomy Plot Analysis: Order-level microbial profiles across four petroleum sludge samples, BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT, and a mixed soil sample, MIX-SOIL

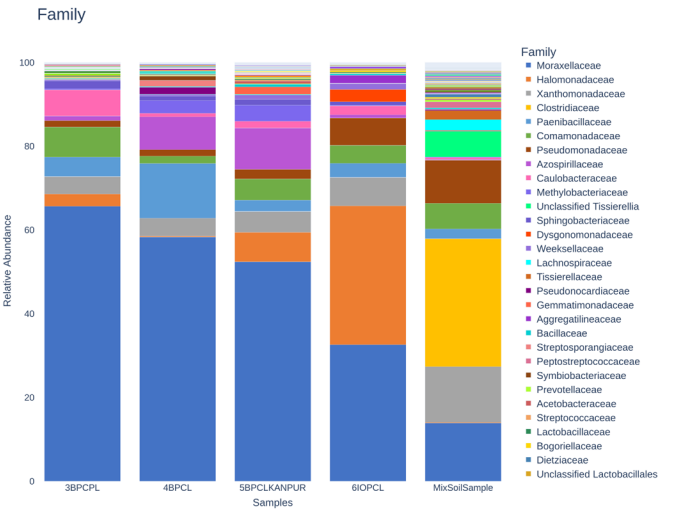

Figure 4. Family-level-Taxonomy Plot Analysis: The families present in the categories of petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT), along with a mixed sample of soil (MIX-SOIL)

samples, Pseudomonadaceae were predominant, comprising 50%-80% of the population, showing high enrichment of hydrocarbon-degrading *Pseudomonas* sp. The slight reduction in Pseudomonadaceae observed in the BPCL-KR and IOCL-MT samples was accompanied by an increase in *Bacillaceae*, *Comamonadaceae*, and Xanthomonadaceae. For MIX-SOIL, families were more evenly represented, with a reduction in Pseudomonadaceae by 10%-15% and a high abundance of *Bacillaceae*, *Lactobacillaceae*, and

Table 2. Numerical alpha-diversity indices calculated from the filtered and rarefied OTU table

Sample	Chao1	Shannon	Simpson	Fisher's alpha
BPC-KD	333	1.73	0.56	31
BPCL-KN	302	1.85	0.64	28
BPCL-KR	341	2.13	0.70	34
IOCL-MT	286	2.10	0.77	25
MIX-SOIL	456	3.08	0.87	61

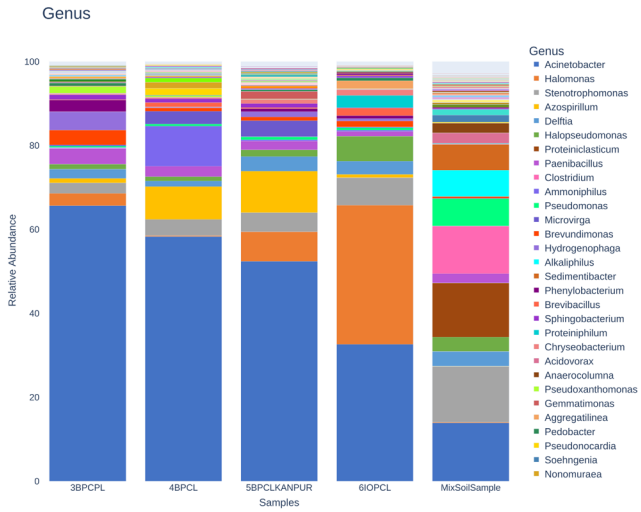

Figure 5. Genus Level-Taxonomy Plot Analysis: Genus-level taxonomic composition of petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT) and the mixed soil sample (MIX-SOIL)

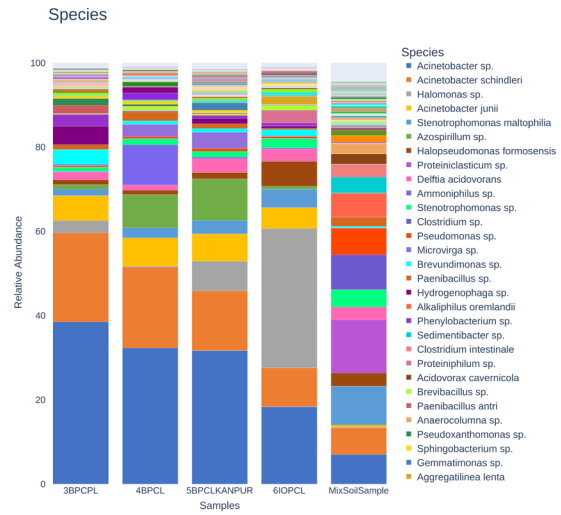

Figure 6. Closest-reference taxonomic annotations from V3-V4 16S rRNA amplicons. Labels below genus level are provisional and should not be interpreted as confirmed species; low-abundance taxa are summarized as others

Streptococcaceae, indicating high diversity. This shows the enrichment of Pseudomonadaceae in the petroleum sludge samples, along with diverse compositions in the mixed soil sample

The genus-level taxonomic composition of the petroleum sludge samples and MIX-SOIL is represented in Figure 5. In the sludge samples, the genus *Pseudomonas* was dominant, owing to the enrichment of hydrocarbon-degrading bacteria. However, in BPCL-KR and IOCL-MT, the reduction in the *Pseudomonas* genus was followed by an increase in other hydrocarbon-related genera, such as *Bacillus, Brevundimonas, Acinetobacter, Hydrogenophaga,* and *Delftia*, indicating site specific microbial adaptability rather than degradative capability. Conversely, MIX-SOIL showed varying genus constituents, with *Pseudomonas* being a minor component, and the presence of soil-related genera such as *Bacillus, Microbacterium, Paenibacillus, Brevibacillus,* and *Phenylobacterium*, indicative of soil microbiota not enriched with pollutants and remaining constant. These results indicate the dominance of the *Pseudomonas* genus in petroleum sites.

The lower-ranked taxonomic annotations returned using the 16S rRNA pipeline are summarised in Figure 6. These labels represent only closest reference assignments, not confirmed

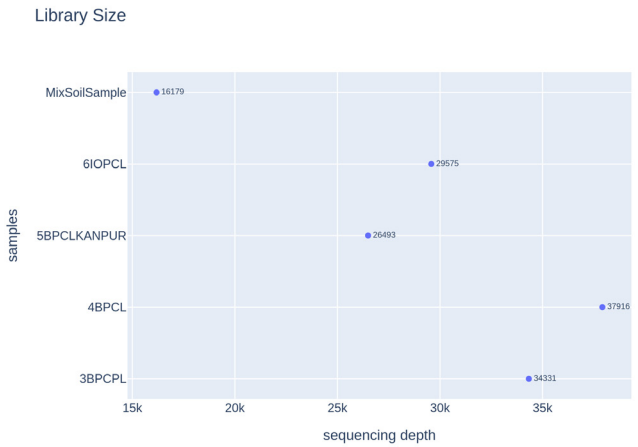

Figure 7. Library Size Overview: Sequencing depth distribution across petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, IOCL-MT) and a mixed soil sample (MIX-SOIL)



Figure 8. Neigh–Curtis PCoA plot showing compositional dissimilarities among the five samples.

species, because the V3-V4 region alone does not provide reliable species-level resolution. Therefore, genus-level patterns, including *Pseudomonas, Acinetobacter, Brevundimonas, Delftia, Hydrogenophaga, Paenibacillus* and *Clostridium*, were emphasized. Low-abundance annotations are grouped as 'Others' where possible to improve readability.

Sequencing result analysis

Figure 7 elucidates the sequencing depth distribution across the petroleum sludge samples (BPC-KD, BPCL-KN, BPCL-KR, and IOCL-MT) and mixed soil sample (MIX-SOIL). All samples accumulated sufficient reads (16,000-38,000) for 16S V3-V4 analysis. MIX-SOIL had the lowest sequencing depth (~16,000 reads). Sequencing depths were higher in BPCL-KN and BPC-KD (~38,000 and ~34,000, respectively). In contrast, sequencing depths were moderate for BPCL-KR and IOCL-MT (~26,000 and ~29,000 reads, respectively). A clear separation can be seen in the figure between the petroleum sludge and soil samples based on sequencing depth, with the sludge samples clustering at higher read counts than MIX-SOIL. The relative spacing among the

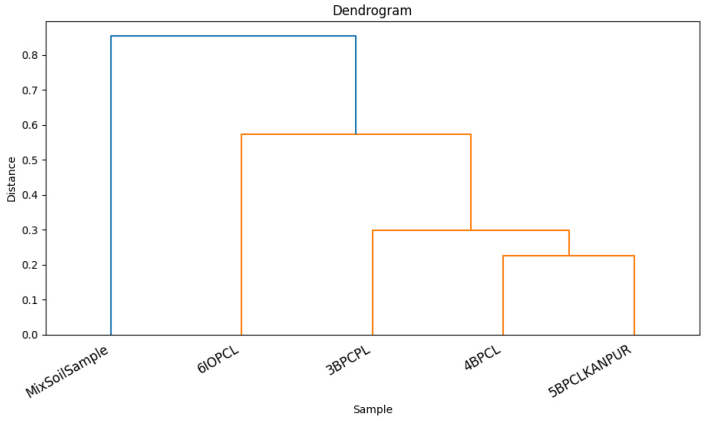

Figure 9. Ward hierarchical-clustering dendrogram based on Bray-Curtis dissimilarity, showing MIX-SOIL as the most compositionally distinct sample

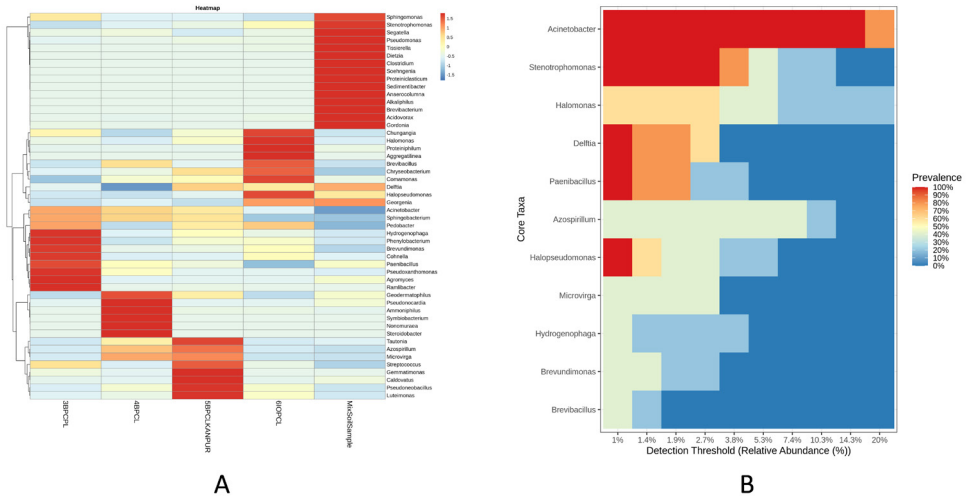

Figure 10. (A) Heatmap Analysis and (B) Core Microbiome

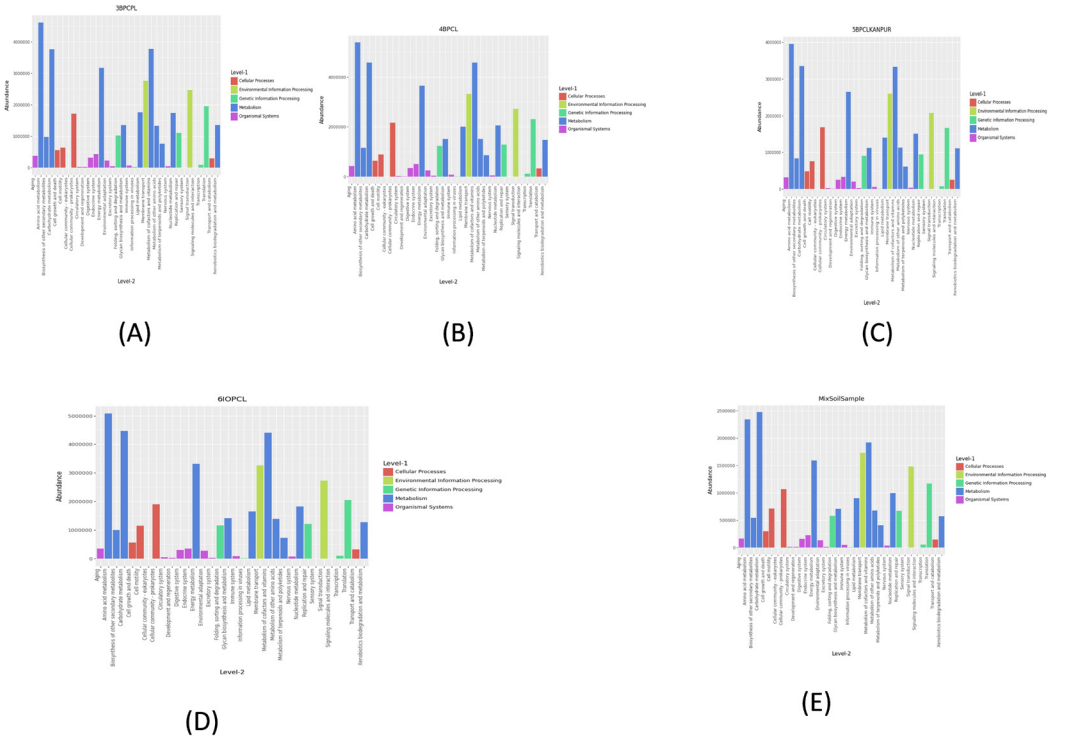

Figure 11. Combined functional prediction plot showing KEGG Level-2 pathway distributions across samples, where (A) represents Metabolism pathways, (B) shows Genetic Information Processing, (C) displays Environmental Information Processing, (D) illustrates Xenobiotic Biodegradation and Metabolism, and (E) depicts Cellular Processes and Community-Level Functions

petroleum sludge samples reflects site-specific differences in sequencing depth.

Alpha diversity

Alpha diversity values were the highest in MIX-SOIL and the lowest in the four petroleum-sludge samples (Table 2). Values were reported numerically rather than qualitatively.

MIX-SOIL had the highest values for Chao1 richness (456), Shannon diversity (3.08), Simpson's diversity (0.87), and Fisher's alpha (61). Among petroleum-sludge samples, Chao1 ranged from 286-341, Shannon from 1.73-2.13, Simpson from 0.56-0.77 and Fisher's alpha from 25-34. These values indicate a richer and more even community in MIX-SOIL and a stronger dominance of fewer taxa in the petroleum sludge samples.

In the petroleum sludge group, IOCL-MT had the lowest Chao1 estimate (286), BPC-KD had the lowest Shannon (1.73) and Simpson (0.56) values, and BPCL-KR had the highest Shannon value (2.13). IOCL-MT shows the highest Simpson's value among the sludge samples (0.77), whereas BPCL-KR had the highest Fisher's alpha (34). These are descriptive comparisons because biological replication was not available.

Culture-based isolation results

Repeated enrichment in crude oil-supplemented Bushnell-Haas medium resulted in growth and morphologically distinct colonies were recovered on Bushnell-Haas agar and purified. However, the supplied dataset did not include isolate counts, colony characteristics, isolate-level 16S sequences, or hydrocarbon degradation assay values. Consequently, isolate-specific quantitative results cannot be reported and must be added to the laboratory records before the final submission.

Beta diversity

Principal Coordinates Analysis (PCoA)

Figure 8 shows the Bray-Curtis PCoA plot, which represents the clustering of the five samples. BPC-KD and BPCL-KN clustered closely, indicating similar hydrocarbon-degrading microbial communities. BPCL-KR and IOCL-MT were moderately spaced apart. MIX-SOIL was positioned farthest from the other samples, reflecting a substantially distinct microbial community composition relative to the petroleum sludge samples. The spatial distribution of the Bray-Curtis PCoA plot highlights greater compositional similarity within the petroleum sludge samples compared to MIX-SOIL.

Hierarchical Clustering (Dendrogram)

Ward clustering based on Bray-Curti's dissimilarity placed BPCL-KN and BPCL-KR closest together (linkage distance of approximately 0.22), followed by BPCL-KD (approximately 0.30). IOCL-MT joined this cluster at approximately 0.57, whereas MIX-SOIL was the most distinct (approximately 0.86; Figure 9).

Heatmap and core microbiome analysis

Figure 10 shows the distribution of microbial genera and the core microbiome within the petroleum sludges and MIX-SOIL. Figure 10A displays the dominant genera within the petroleum sludge: *Pseudomonas, Acinetobacter, Rhodococcus, Brevundimonas,* and *Hydrogenophaga,* all known to have hydrocarbon-degrading potential. These are much more abundant in petroleum sludge. MIX-SOIL, however, contains microorganisms belonging to the genera *Dietzia, Streptomyces, Paenibacillus,* and *Bacillus,* as found within normal soil, not under the selective pressure due to hydrocarbon presence.

Figure 10B, or the core microbiome, shows four genera (*Pseudomonas, Acinetobacter, Rhodococcus,* and *Bacillus*) that are crucial in hydrocarbon degradation, signifying their ecological resilience and significance in hydrocarbon degradation in soil ecosystems over time. Although hydrocarbon degraders are mostly specialists in contaminated sites, the presence of ecologically stable and diverse hydrocarbon degrading ecologies in non-contaminated soil sites, as observed in MIX-SOIL, provides an important indicator of their ability to be resilient in soil ecosystems over a period of time and their functional significance in their degradation process in soil ecosystems at a microbial level.

Functional analysis

Functional prediction analysis with PICRUSt2 in Figure 11 showed the enrichment of petroleum-sludge samples with major KEGG functional groups, reflecting the ability of hydrocarbon-degrading communities to adapt to their metabolic requirements (Figure 11A-E). In Figure 11A, the function Metabolism reflects the enhanced activity of carbohydrate, amino acid, lipid, and energy metabolic pathways; this function is integrated into the figure because the ability of hydrocarbon-degrading bacteria to remain versatile in their metabolism is an added advantage in the process of hydrocarbon degradation. Similarly, in Figure 11B, Genetic Information Processing, the enhanced activity of DNA replication and repair, transcription, and translation is indicated; this functional group is integrated into the figure because hydrocarbon degradation communities had enhanced activity corresponding to high stress levels from petroleum in their environments. Similarly, in Figure 11C, Environmental Information Processing, enhanced activity is observed in membrane transporters, two-component systems, and ABC transport systems; the functional group is integrated into the figure because hydrocarbon degradation communities had enhanced activity attributable to increased nutrient uptake, degradation, and petroleum-induced stress.

Figure 11D (Xenobiotic Degradation) clearly shows the overrepresentation of alkane monooxygenase (alkB)-, catechol dioxygenase (catA)-, and naphthalene dioxygenase (nahA)-related functions in the sludge samples, indicating the presence of specialized degraders for petroleum hydrocarbon degradation. Figure 11E (Cellular Processes and Community Functions) focuses on the balanced participation of secondary metabolite biosynthesis, nitrogen-related processes, quorum sensing, and soil nutrient cycles for MIX-SOIL, indicating the natural complex environmental processes occurring in the soil

DISCUSSION

This study characterized the bacterial communities in petroleum sludge from four industrial locations in Uttar Pradesh, India. The amplicon profiles were dominated by *Pseudomonadota*, particularly the genus *Pseudomonas*. Other genera with documented hydrocarbon degradation potential, including *Acinetobacter, Rhodococcus, Brevundimonas, Hydrogenophaga* and *Bacillus*, were also detected. PICRUSt2 suggested enrichment of functions associated with alkB, nahA, and catA; however, these were predictions based on 16S rRNA profiles and did not demonstrate gene abundance, expression, or enzyme activity.

The numerical alpha-diversity metrics showed lower richness and evenness in the petroleum sludge than in the MIX-SOIL. The Bray-Curtis ordination and dendrogram also separated MIX-SOIL from the sludge samples. As the design included only one composite sample per location, these patterns were descriptive and should not be interpreted as statistically significant site effects.

The dominance of *Pseudomonas* species in petroleum-contaminated sites noted in this study is supported by other metagenomic studies worldwide. For example, Li et al. indicated that petroleum contamination increases the abundance of *Pseudomonas, Mycobacterium,* and *Pseudoxanthomonas,* which produce enzymes that degrade toluene, xylene, and polycyclic aromatic compounds (PAHs), respectively.[31] Moreover, Pacwa-Plociniczak et al. indicated that *Rhodococcus* and *Pseudomonas* species were found in old petroleum-contaminated soils that enhanced xenobiotic biometabolism and petroleum biodepletion, indicating that these species play a role in petroleum biodegradation.[32]

The present results agree with Jurelevicius et al., who observed lower microbial diversity and enhanced dominance of hydrocarbon-disposing families, such as *Halomonadaceae* and *Bacillaceae* with increasing hydrocarbon contamination in Antarctic soil samples.[33] Basim et al. demonstrated analogous patterns in petroleum-contaminated soils with diminished microbial diversity and increased dominance of hydrocarbon-disposing families.[34] Consistent with these observations, our results from petroleum-contaminated sites indicate that hydrocarbon-mediated microbial selection pressures result in specific microbial communities.

The presence of *Pseudomonas*, *Acinetobacter,* and *Rhodococcus* at petroleum-polluted sites also verified their documented applicability to hydrocarbon biodegradation.[35] Chicca et al. documented the diesel- and PAH-degrading abilities of *Pseudomonas, Arthrobacter,* and *Microbacterium* species, validating their applicability in petroleum hydrocarbon biodegradation.[36] Aboud et al. found that *Pseudomonas stutzeri, Bacillus cereus,* and *Psychrobacter* species have high diesel-degrading abilities in oil sludge, thus validating their adaptability to hydrocarbon degradation.[37]

Our functional prediction analysis high enrichment of the following functions: xenobiotic degradation, lipid oxidation, and energy-related functions.[38,39] This is consistent with the global metagenomic analysis. Various studies, including those by Eze, have isolated different bacterial consortia, including the genera *Acidocella* and *Pseudomonas*, which participate in the aromatic ring cleavage of BTEX.[40] This indicates that these genera remain part of an unchanged network involved in the degradation of hydrocarbons, where Pseudomonadaceae is the major hydrocarbon-degrading agent.

Our results are consistent with the reported importance of plasmid-mediated degradation and adaptive enzymatic approaches, as revealed by Mirdamadian et al.[41] and Goma-Tchimbakala EJC et al.[42] The results confirmed the significance of gene transfer and redundancy in the establishment of effective hydrocarbon-degrading consortia that can sustain themselves according to pollution scenarios.

The present study also verified the universality of petroleum-polluted environments, placing strong selective pressure on microbes, which triggers the dominance of metabolically flexible species such as *Pseudomonas, Acinetobacter, Rhodococcus,* and *Bacillus*. The aforementioned bacteria possess genes for hydrocarbon degradation and show strong adaptability to environmental factors through biosurfactant production and use of hydrocarbons as carbon sources.

Through the synergy of culture-based isolation techniques and metagenomic approaches, we identified important microbes and enzymatic processes that contribute to hydrocarbon biodegradation. Our results provide insights for developing sustainable bioremediation microbial cultures. Future research should improve understanding of the regulatory processes in enzymatic biodegradation and assess applications of native microbial communities in bioaugmentation or biostimulation.

CONCLUSION

Petroleum-sludge communities from the sampled industrial sites in Uttar Pradesh were dominated by *Pseudomonadota* and the genera *Pseudomonas*, *Acinetobacter, Rhodococcus* and *Bacillus.* MIX-SOIL had a higher numerical alpha diversity than the sludge samples, consistent with selective enrichment under hydrocarbon exposure. Predicted xenobiotic degradation functions, including alkB-, nahA-, and catA-associated pathways, identify testable targets but require confirmation by targeted PCR, transcript analysis, enzyme assays, or shotgun metagenomics. The culture enrichment and amplicon results provide a basis for future replicated studies and isolate-level degradation experiments.

ACKNOWLEDGMENTS

The authors would like to acknowledge Amity University, Lucknow, for providing the required resources for the research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORS' CONTRIBUTION

MD conceptualized and designed the study. HY performed the experiments. MD and HY analysed the data and wrote the manuscript. MD, HY and MHS reviewed and revised the manuscript. All authors read and approved the final manuscript for publication.

FUNDING

None.

DATA AVAILABILITY

The datasets generated and/or analysed during the current study are available in the NCBI database repository, Accession number: [SUB15270059]. SRA: Processed (5 objects): SRR33230738, SRR33230737, SRR33230736, SRR33230735, SRR33230734.

ETHICS STATEMENT

Not applicable.

REFERENCES

1. Mohammadi L, Rahdar A, Bazrafshan E, Dahmardeh H, Susan MA, Kyzas GZ. Petroleum hydrocarbon removal from wastewaters: a review. *Processes*. 2020;8(4):447. doi: 10.3390/pr8040447
2. Mohanta S, Pradhan B, Behera ID. Impact and remediation of petroleum hydrocarbon pollutants on agricultural land: a review. *Geomicrobiol J.* 2024;41(4):345-359. doi: 10.1080/01490451.2023.2243925
3. Patel AB, Shaikh S, Jain KR, Desai C, Madamwar D. Polycyclic aromatic hydrocarbons: sources, toxicity, and remediation approaches. *Front Microbiol*. 2020;11:562813. doi: 10.3389/fmicb.2020.562813
4. Allison E, Mandler B. Petroleum and the Environment. *American Geosciences Institute.* 2018
5. Mai Y, Wang Y, Geng T, et al. A systematic toxicologic study of polycyclic aromatic hydrocarbons on aquatic organisms via food-web bioaccumulation. *Sci Total Environ.* 2024;929:172362. doi: 10.1016/j.scitotenv.2024.172362
6. Tarrad MR, Ali HA, Salih BH, Obayes ZM, Jawad RM, Jawad AM. Polycyclic Aromatic Hydrocarbon (PAHs) Pollutants: Toxic Organic Compounds, Hydrocarbons and Indicators of Exposure To Xenobiotic. *Current Clinical and Medical Education.* 2024;2(09):19-35.
7. Onyena AP, Nkwoji JA, Chukwu LO, Walker TR, Sam K. Risk assessment of sediment PAH, BTEX, and emerging contaminants in Chanomi Creek Niger Delta, Nigeria. *Environ Monit Assess*. 2023;195(9):1080. doi: 10.1007/s10661-023-11703-x
8. Allamin IA, Shukor MY. Phytoremediation of PAHs in contaminated soils: a review. *Bioremed Sci Technol Res.* 2021;9(2):1-6. doi: 10.54987/bstr.v9i2.609
9. Gupte A, Tripathi A, Patel H, Rudakiya D, Gupte S. Bioremediation of polycyclic aromatic hydrocarbon (PAHs): a perspective. *Open Biotechnol J*. 2016;10(1):363. doi: 10.2174/1874070701610010363
10. Jing L, Li P, Li Z, Ma D, Hu J. Influence of π-π interactions on organic photocatalytic materials and their performance. *Chem Soc Rev.* 2025;54(4):2054-2090. doi: 10.1039/D4CS00029C
11. Wu B, Rayaroth M, Aeppli C. Hydrophobic model systems for oil film photooxidation: part II: direct and indirect pathways controlled by oxygen and molecular structure. *Environ Sci Process Impacts*. 2026;28(7):1974-1986. doi: 10.1039/d6em00173d
12. Khurana P, Das RK, Brar SK. From prescription to

pollution: environmental behavior and breakdown of fluoxetine. *Environ Sci: Water Res Techno.* 2025;11(12):2825-2843 doi: 10.1039/D5EW00636H

13. Liu J, Zhang AN, Liu YJ, Liu Z, Liu Y, Wu XJ. Analysis of the mechanism for enhanced pyrene biodegradation based on the interactions between iron-ions and *Rhodococcus ruber* strain L9. *Ecotoxicol Environ Saf.* 2021;225:112789. doi: 10.1016/j.ecoenv.2021.112789
14. Banerjee S, Gupta N, Pramanik K, et al. Microbes and microbial strategies in carcinogenic polycyclic aromatic hydrocarbons remediation: a systematic review. *Environ Sci Poll Res.* 2024;31(2):1811-1840. doi: 10.1007/s11356-023-31140-0
15. Perez-Garcia P, Sass K, Wongwattanarat S, et al. Microbial plastic degradation: enzymes, pathways, challenges, and perspectives. *Microbiol Mol Biol Rev.* 2025;89(4):e00087-24. doi: 10.1128/mmbr.00087-24
16. Gupta R, Gupta N. Lipid biosynthesis and degradation. In: *Fundamentals of Bacterial Physiology and Metabolism*. Springer; 2021. doi: 10.1007/978-981-16-0723-3_18
17. Mallick S, Chakraborty J, Dutta TK. Role of oxygenases in guiding diverse metabolic pathways in the bacterial degradation of low-molecular-weight polycyclic aromatic hydrocarbons: a review. *Crit Rev Microbiol.* 2011;37(1):64-90. doi: 10.3109/1040841X.2010.512268
18. Zhang T, Zhang H. Microbial consortia are needed to degrade soil pollutants. *Microorganisms*. 2022;10(2):261. doi: 10.3390/microorganisms10020261
19. Yesankar PJ, Pal M, Patil A, Qureshi A. Microbial exopolymeric substances and biosurfactants as 'bioavailability enhancers' for polycyclic aromatic hydrocarbons biodegradation. *Int J Environ Sci Technol*. 2023;20(5):5823-5844. doi: 10.1007/s13762-022-04068-0
20. Kaczorek E, Pacholak A, Zdarta A, Smulek W. The impact of biosurfactants on microbial cell properties leading to hydrocarbon bioavailability increase. *Colloids Interfaces.* 2018;2(3):35. doi: 10.3390/colloids2030035
21. Vidal-Silva IM, Loza A, Gutierrez-Rios RM. Unlocking microbial potential: advances in omics and bioinformatics for aromatic hydrocarbon degradation. *World J Microbiol Biotechnol.* 2025;41(10):384. doi: 10.1007/s11274-025-04592-4
22. Sokal S, Palsania P, Kaushik G. Bioremediation and functional metagenomics: advances, challenges, and opportunities. In: Kumar V, Thakur IS, eds. *Omics Insights in Environmental Bioremediation*. Springer; 2022. doi: 10.1007/978-981-19-4320-1_1
23. Mukherjee A, Chettri B, Langpoklakpam JS, et al. Bioinformatic approaches including predictive metagenomic profiling reveal characteristics of bacterial response to petroleum hydrocarbon contamination in diverse environments. *Sci Rep*. 2017;7(1):1108. doi: 10.1038/s41598-017-01126-3
24. Junkins EN, Stevenson BS. Using plate-wash PCR and high-throughput sequencing to measure cultivated diversity for natural product discovery efforts. *Front Microbiol.* 2021;12:675798. doi: 10.3389/fmicb.2021.675798
25. Karthikeyan S, Hatt JK, Kim M, Spain JC, Huettel M, Kostka JE, Konstantinidis KT. A novel, divergent alkane monooxygenase (alkB) clade involved in crude oil biodegradation. *Environ Microbiol Rep*. 2021;13(6):830-840. doi: 10.1111/1758-2229.13018
26. Abbasian F, Palanisami T, Megharaj M, Naidu R, Lockington R, Ramadass K. Microbial diversity and hydrocarbon degrading gene capacity of a crude oil field soil as determined by metagenomics analysis. *Biotechnol Prog.* 2016;32(3):638-648. doi: 10.1002/btpr.2249
27. Kadnikov VV, Ravin NV, Sokolova DS, et al. Metagenomic and culture-based analyses of microbial communities from petroleum reservoirs with high-salinity formation water, and their biotechnological potential. *Biology.* 2023;12(10):1300. doi: 10.3390/biology12101300
28. Alkaabi RS. *Meta-analysis of microbial communities in worldwide petroleum fields: an international analysis between 16S and shotgun metagenomic datasets* [master's thesis]. Khalifa University; 2021
29. Nazina TN, Sokolova DS, Babich TL, et al. Phylogenetic diversity of microorganisms from the sludge of a biogas reactor processing oil-containing and municipal waste. *Microbiology.* 2018;87(3):416-424. doi: 10.1134/S0026261718030074
30. Sierra-Garcia IN, Dellagnezze BM, Santos VP, et al. Microbial diversity in degraded and non-degraded petroleum samples and comparison across oil reservoirs at local and global scales. *Extremophiles*.2017;21(1):211-229. doi: 10.1007/s00792-016-0897-8
31. Li YQ, Xin Y, Li C, Liu J, Huang T. Metagenomics-metabolomics analysis of microbial function and metabolism in petroleum-contaminated soil. *Braz J Microbiol.*2023;54(2):935-947. doi: 10.1007/s42770-023-01000-7
32. Pacwa-Płociniczak M, Biniecka P, Bondarczuk K, Piotrowska-Seget Z. Metagenomic functional profiling reveals differences in bacterial composition and function during bioaugmentation of aged petroleum-contaminated soil. *Front Microbiol.* 2020;11:2106. doi: 10.3389/fmicb.2020.02106
33. Jurelevicius D, Pereira RS, da Mota FF, et al. Metagenomic analysis of microbial communities across a transect from low to highly hydrocarbon-contaminated soils in King George Island, Maritime Antarctica. *Geobiology.* 2021;20(1):98-111. doi: 10.1111/gbi.12472
34. Basim Y, Mohebali G, Jorfi S, et al. Biodegradation of total petroleum hydrocarbons in contaminated soils using indigenous bacterial consortium. *Environ Health Eng Manag.* 2020;7(2):127-133. doi: 10.34172/EHEM.2020.15
35. Wang M, Garrido-Sanz D, Sansegundo-Lobato P, et al. Soil microbiome structure and function in ecopiles used to remediate petroleum-contaminated soil. *Front Environ Sci*. 2021;9:624070. doi: 10.3389/fenvs.2021.624070
36. Chicca I, Becarelli S, Bernabei G, Siracusa G, Di Gregorio S. Innovative culturomic approaches and predictive functional metagenomic analysis: the isolation of hydrocarbonoclastic bacteria with plant growth

promoting capacity. *Water.* 2022;14(2):142. doi: 10.3390/w14020142

37. Aboud EM, Burghal AA, Laftah AH. Genetic identification of hydrocarbons degrading bacteria isolated from oily sludge and petroleum-contaminated soil in Basrah City, Iraq. *Biodiversitas.* 2021;22(4):1934-1939. doi: 10.13057/biodiv/d220441
38. Nyarko HD, Okpokwasili GC, Joel OF, Galyuon IAK. Isolation and characterization of carbonoclastic bacteria diversity in oil-contaminated soil in Cape Coast Metropolis, Ghana. *Microbiol Res J Int.* 2020;30(1):62-74. doi: 10.9734/mrji/2020/v30i130191
39. Olukunle O. Phylogenetic analysis of hydrocarbon degrading bacteria associated with crude oil polluted soil from Mesogar community, Delta State, Nigeria. *Niger J Biotechnol.* 2020;36(2):62-76. doi: 10.4314/njb.v36i2.8
40. Eze MO. Metagenome analysis of a hydrocarbon-degrading bacterial consortium reveals the specific roles of BTEX biodegraders. *Genes*. 2021;12(1):98. doi: 10.3390/genes12010098
41. Mirdamadian SH, Emtiazi G, Golabi M, Ghanavati H. Biodegradation of petroleum and aromatic hydrocarbons by bacteria isolated from petroleum-contaminated soil. *J Pet Environ Biotechnol.* 2010;1(1):1-5. doi: 10.4172/2157-7463.1000102
42. Goma-Tchimbakala EJC, Pietrini I, Goma-Tchimbakala J, Corgnati SP. Use of shotgun metagenomics to assess the microbial diversity and hydrocarbons degrading functions of auto-mechanic workshop soils polluted with gasoline and diesel fuel. *Microorganisms.* 2023;11(3):722. doi: 10.3390/microorganisms11030722