

Metagenomic Approaches for Studying Microbial Community of Industrial Effluent Contaminated Soil Across Dadra and Nagar Haveli

Nipun Bariya^{1*,2}, Narayan Kamath², Aparna Tailor³, Mayuri Dholaria¹

¹Naran Lala College of Professional and Applied Sciences, Navsari, Gujarat, India.

²Department of Microbiology, NAMO MERI Silvassa, Dadra and Nagar Haveli-396230, India

³Department of Microbiology, Dr. APJ Abdul Kalam Government College, Silvassa, Dadra and Nagar Haveli – 396230, India

*Corresponding Author: nipun.bariya015@gmail.com

*ORCID: 0009-0001-4956-5497

DOI: <https://doi.org/10.63001/tbs.2026.v21.i04.pp258-276>

KEYWORDS

Metagenomics;
Industrial Effluent;
Contaminated Soil;
Dadra and Nagar Haveli;
Silvassa; Microbial Community;
16S rRNA; Bioremediation;
Heavy Metals;
Functional Annotation;
Xenobiotics; KEGG; COG

Abstract

Dadra and Nagar Haveli (DNH), a Union Territory of western India, hosts one of the nation's most densely industrialized landscapes, with over 5,100 large, medium, and small-scale industries operating along the ecologically sensitive Daman Ganga River corridor. Decades of inadequately treated industrial effluent discharge from textile dyeing, electroplating, tanning, pharmaceutical, and chemical manufacturing units have profoundly altered soil ecosystems across the Silvassa industrial belt, introducing complex mixtures of heavy metals, petroleum hydrocarbons, xenobiotic organic compounds, and pharmaceutical active compounds into the soil matrix. Metagenomic studies of industrial effluent-contaminated soils in Silvassa consistently document communities dominated by Pseudomonadota (Proteobacteria), Actinomycetota, Bacillota, Myxococcota, Acidobacteriota, and Thermodesulfobacteriota. Key genera—*Pseudomonas*, *Bacillus*, *Cupriavidus*, *Novosphingobium*, *Sulfuricurvum*, *Variovorax*, and *Clostridium*—encode hydrocarbon-degrading enzymes (ring-hydroxylating dioxygenases, monooxygenases, dehalogenases) and heavy-metal resistance determinants (*czc*, *chr*, *ars* operons). KEGG and COG functional annotation reveals active metabolic pathways for benzoate, PAH, BTEX, chloroalkane, chloroalkene, and nitrotoluene catabolism converging on the TCA cycle. Co-selection of antibiotic resistance genes (ARGs) via mobile genetic elements in metal-contaminated soils constitutes an emerging One Health concern. Industrial effluent-contaminated soils of Dadra and Nagar Haveli harbor a taxonomically diverse and functionally rich indigenous microbiome with considerable untapped bioremediation potential. Metagenomic science provides the resolution needed to translate community genomic data into targeted bio-stimulation, bioaugmentation, and phytoremediation strategies. Longitudinal, multi-omics monitoring programs and integration of long-read sequencing platforms are prioritized as future research directives for this critically understudied region.

Article Info

Received: 25.09.2026

Accepted: 11.10.2026

Published: 23.10.2026

1. INTRODUCTION

Soil represents one of the most biologically diverse ecosystems on Earth, harboring an estimated 10^8 – 10^9 microbial cells and thousands of distinct taxa per gram, whose collective metabolic activities sustain the biogeochemical cycles upon which all terrestrial life depends. In industrialized regions, the uncontrolled discharge of untreated or insufficiently treated effluents from manufacturing industries exerts severe selective pressures on these soil communities, fundamentally restructuring their taxonomic composition, suppressing sensitive taxa,

enriching pollution-adapted lineages, and degrading essential ecosystem functions including nutrient cycling, organic matter mineralization, and soil structural maintenance (Mi et al., 2025; C. Su et al., 2023).

Dadra and Nagar Haveli (DNH), a Union Territory flanked by Gujarat and Maharashtra in western India, epitomizes this challenge. The territory's administrative capital, Silvassa, and its surrounding industrial estates—including Rakholi, Surangi, Kherdi, Dapada, and Saily—constitute one



of India's fastest-growing and most densely concentrated industrial clusters. More than 5,100 registered industrial units, spanning textile dyeing and sizing, plastic and polymer manufacturing, pharmaceutical production, electroplating and surface finishing, chemical synthesis, food processing, and engineering fabrication, discharge large volumes of effluents into the Daman Ganga River basin and surrounding land (Bariya et al., 2025; Nipun Mayuri Dholaria, 2024). The Daman Ganga River, which traverses 131.30 km from its origin in the Sahyadri Hills through Maharashtra, Gujarat, and DNH before draining into the Arabian Sea, receives cumulative industrial and domestic waste loads from the Silvassa, Vapi, and Daman industrial corridors, exposing adjacent soils to persistent multi-pollutant contamination.

Characterizing the microbial ecology of these contaminated soils has historically been constrained by the limitations of culture-dependent methodologies. Fewer than 1% of environmental microorganisms are cultivable under standard laboratory conditions, meaning that conventional culture-based surveys provide a profoundly incomplete picture of in situ community diversity and function (Hugenholtz, 2002). The emergence of culture-independent, DNA sequence-based metagenomics over the past two decades has transformed environmental microbiology, enabling direct, comprehensive access to the total genomic content of complex microbial assemblages without prior cultivation. Metagenomics encompasses two principal methodological pillars: targeted amplicon sequencing of phylogenetic marker genes (principally the 16S rRNA gene for bacteria and archaea, and the ITS region for fungi), and whole-genome shotgun (WGS) sequencing of total community DNA to simultaneously resolve taxonomic identity and functional gene content (Nwachukwu & Babalola, 2022; Sousa et al., 2022).

Recent primary investigations in Silvassa mark a scientific watershed for DNH soil microbiology. Nipun *et al.* (2024) conducted the first dedicated metagenomic exploration of industrial effluent-contaminated soils from Rakholi, Surangi, and Kherdi in Silvassa, documenting community dominance by Pseudomonadota, Actinomycetota, Bacillota, Myxococcota, Acidobacteriota, and

Thermodesulfobacteriota, with pollutant-degrading genera such as *Pseudomonas*, *Bacillus*, *Clostridium*, *Candidatus Entotheonella*, and *Desulforomonas* indicating significant in situ bioremediation capacity. Building on this foundation, Bariya *et al.* (2025) extended the metagenomic investigation to effluent-contaminated soils from Hindustan Unilever Ltd. (Dapada), and two additional Silvassa-area sites, employing KEGG and COG functional annotation to reveal a comprehensive aromatic hydrocarbon degradation network encompassing benzoate, PAH, toluene, xylene, ethylbenzene, nitrotoluene, chloroalkane, and chloroalkene catabolism. Analogous investigations from common effluent treatment plants (CETPs) at Baddi Vasudeva *et al.* (2022) and contaminated soils near Bhilai steel plant Sandhu *et al.* (2022) provide essential comparative context for interpreting DNH findings within the broader landscape of Indian industrial soil metagenomics.

Despite this emerging body of evidence, the metagenomic characterization of industrial effluent-contaminated soils in DNH remains in early stages relative to the scale and complexity of the contamination challenge. Critical knowledge gaps persist regarding: the spatial distribution of contamination-driven microbial community shifts across different industrial estate types; the identity and metabolic activity of archaeal, fungal, and viral community members; the dynamics of antibiotic resistance gene (ARG) co-selection in metal-polluted soils; and the translational pipeline from metagenomic discovery to operational bioremediation strategy. This review synthesizes the current state of knowledge, critically examines methodological advances and limitations, and delineates a research agenda to address these gaps in the specific context of DNH.

2. INDUSTRIAL EFFLUENT POLLUTION IN DADRA AND NAGAR HAVELI

2.1 Industrial Profile of the Union Territory

Dadra and Nagar Haveli (DNH) covers an area of 491 km² (latitude 20°10' N – 20°25' N; longitude 72°50' E – 73°15' E) and is home to one of India's most rapidly industrialized Special Economic Zones. The industrial landscape is anchored by several major estate clusters: Rakholi Industrial Estate, Kherdi Industrial Estate, Surangi Industrial

Estate, Dapada Industrial Area, Saily Industrial Area, and Masat Industrial Area. The Central Pollution Control Board (CPCB) has identified polluted river stretches in the Daman Ganga basin attributable to this industrial density (Indiastat, 2022). The territory's economic model—characterized by low land costs, proximity to the chemical and pharmaceutical hub of Vapi (Gujarat), good port access through Daman, and a business-friendly regulatory environment—has attracted a diverse industrial mix whose combined effluent profile is chemically heterogeneous and ecotoxicologically significant.

The principal industrial sectors and their associated effluent contaminant profiles are as follows: (i) Textile dyeing and sizing units contribute synthetic azo dyes, reactive dyes, dispersants, surfactants, and auxiliary chemicals characterized by high biological oxygen demand (BOD), chemical oxygen demand (COD), and total dissolved solids (TDS); (ii) Electroplating and surface finishing operations discharge heavy metals—chromium (both hexavalent Cr^{6+} and trivalent Cr^{3+}), cadmium (Cd), nickel (Ni), copper (Cu), and zinc (Zn)—along with cyanide compounds; (iii) Pharmaceutical and active pharmaceutical ingredient (API) manufacturing generates API residues, solvents (dichloromethane, acetonitrile, methanol), and biological growth media; (iv) Plastic and polymer processing discharges phthalate plasticizers, bisphenol A, and polycyclic aromatic hydrocarbons (PAHs); (v) Food and edible oil processing contributes high-BOD organic loading; and (vi) Engineering fabrication and automotive units release petroleum hydrocarbons, lubricants, and hydraulic fluids (Bariya et al., 2025; Nipun et al., 2024; Patel et al., 2016; Thomas et al., 2020).

2.2 Pathways of Soil Contamination

Industrial effluents reach DNH soils through multiple pathways: (1) direct discharge of untreated or inadequately treated effluents onto adjacent land, either through open drains or unauthorized dumping; (2) seepage from unlined effluent storage lagoons and treatment ponds; (3) flood-driven lateral transport during the intense monsoon season (June–September) when the Daman Ganga and its tributaries overflow their banks; (4) irrigation of agricultural lands with contaminated river water; (5)

aerial deposition of particulate-bound heavy metals from industrial stacks; and (6) leachate percolation from industrial solid waste disposal sites (Feng et al., 2018; Sandhu et al., 2022; Sandhu & Paul, 2021). The confluence of these pathways produces a spatially heterogeneous contamination landscape in which proximity to industrial estates, soil texture, pH, and organic matter content interact to determine the speciation, bioavailability, and persistence of individual contaminants.

Heavy metals—particularly Cr, Pb, Cd, Ni, Cu, and Zn—are of primary concern because they are non-biodegradable, persist indefinitely in soil matrices, and accumulate through bioconcentration pathways. Studies from the analogous Vapi industrial cluster (classified among India's five most polluted industrial areas by the Central Pollution Control Board) document soil heavy metal concentrations exceeding Indian regulatory standards by factors of 2–20 for chromium, lead, and cadmium. Given DNH's ecological and hydrological connectivity with the Vapi cluster through the shared Daman Ganga–Kolak river system, comparable soil metal burdens are plausible across DNH's industrial zones (Mi et al., 2025; H. Su et al., 2022). Concurrent with metal contamination, petroleum hydrocarbon and PAH loading from vehicle maintenance operations and chemical storage facilities introduces persistent organic pollutants (POPs) whose biodegradation by indigenous soil microorganisms is of direct practical relevance to remediation planning (Jiang et al., 2018; Muccee & Ejaz, 2020; Sandhu et al., 2022).

2.3 Physio-Chemical Characterization of Contaminated Soils

Bariya *et al.* (2025) conducted comprehensive physio-chemical characterization of effluent-contaminated soil samples collected in August 2023 from three Silvassa-area locations (SA4: Hindustan Unilever Ltd., Dapada, 20.183°N, 73.030°E; SA5 and SA6 from additional industrial sites). Analysis revealed the presence of macro- and micro-elements including metals and non-metals, alongside various petroleum and aromatic hydrocarbons. Nipun *et al.* (2024) similarly documented elevated concentrations of polyaromatic compounds and heavy metals in soils from Rakholi, Surangi, and Kherdi. These physio-chemical

parameters—pH, organic carbon, nitrogen, electrical conductivity, and metal concentrations—function as primary predictors of microbial community composition, diversity, and functional capacity in contaminated soils (Feng et al., 2018; Mi et al., 2025).

3. METAGENOMIC METHODOLOGICAL FRAMEWORK

3.1 Environmental DNA Extraction from Industrial Effluent-Contaminated Soils

The quality and representativeness of metagenomic data are fundamentally determined by the integrity of environmental DNA (eDNA) extraction from the soil matrix. Industrial effluent-contaminated soils present particularly challenging extraction conditions due to the abundant presence of co-extracted PCR inhibitors: humic acids and fulvic acids that complex with DNA; heavy metal ions that denature DNA polymerases; residual industrial surfactants that disrupt cell lysis efficiencies; and organic solvent residues that interfere with column-based purification chemistries. In Silvassa-area studies, Nipun *et al.* (2024) and Bariya *et al.* (2025) employed total soil DNA extraction followed by whole genome sequencing library preparation, with library quantification and validation using Qubit fluorometric quantification and the Agilent 4150 TapeStation system prior to sequencing.

Recommended extraction protocols for DNH-type contaminated soils incorporate a bead-beating step (5.5 m/s, 30 seconds, using 0.1 mm silica-zirconia beads) for mechanical cell disruption, followed by polyvinylpyrrolidone (PVPP) treatment to chelate humic acids, and column-based DNA purification (DNeasy PowerSoil Pro, Qiagen, or equivalent). Yield thresholds of ≥ 50 ng/ μ L and DNA integrity number (DIN) ≥ 7 , assessed by TapeStation, are appropriate quality benchmarks for whole-genome shotgun library preparation. For 16S rRNA amplicon libraries, lower-purity DNA may be acceptable provided PCR amplification efficiency is validated by comparison of amplicon band intensities across samples (Rekadwad et al., 2024; Vasudeva et al., 2022).

3.2 16S rRNA Amplicon Sequencing

The 16S ribosomal RNA gene, universally conserved across prokaryotes while containing nine hypervariable regions (V1–V9) of discriminatory

phylogenetic value, remains the standard molecular marker for bacterial and archaeal community profiling in environmental samples. For Illumina MiSeq paired-end sequencing (2 $\times$ 250 bp or 2 $\times$ 300 bp), the V3–V4 hypervariable region combination (targeted by primers 341F/806R) provides the optimal balance of taxonomic resolution and amplification universality across divergent environmental taxa. Amplicons are sequenced at a depth of $\geq 10,000$ reads per sample to adequately capture community diversity, and data are processed through QIIME2 (Bolyen et al., 2018) or MOTHUR pipelines to generate amplicon sequence variants (ASVs) via the DADA2 denoising algorithm, with taxonomic classification against the SILVA v138 database at $\geq 97\%$ sequence identity (Ahmad et al., 2021; Bolyen et al., 2018; Mishra et al., 2025; Rekadwad et al., 2024).

Diversity analyses include alpha diversity (species richness: Chao1, ACE; diversity: Shannon entropy, Simpson index; phylogenetic diversity: Faith's PD) and beta diversity (Bray–Curtis dissimilarity, weighted and unweighted UniFrac distances) visualized by principal coordinates analysis (PCoA). Permutational multivariate analysis of variance (PERMANOVA) is applied to test the statistical significance of community composition differences between contaminated and reference soils, and between sites with different industrial effluent types. Indicator species analysis (ISA) identifies taxa significantly associated with specific contamination profiles (Allamin et al., 2020; Mishra et al., 2025; Nipun et al., 2024).

3.3 Whole-Genome Shotgun (WGS) Metagenomics

WGS metagenomics provides simultaneous resolution of taxonomic profiles and functional gene content from a single sequencing library, making it the methodology of choice for investigating the bioremediation potential of industrial effluent-contaminated soils. Bariya et al. (2025) employed WGS metagenomics to analyze Silvassa contaminated soils, generating comprehensive functional annotation data through KEGG and COG databases. Sandhu et al. (2022), investigating PAH- and PCB-contaminated soils near the Bhilai steel plant using Oxford Nanopore GridION Technology, achieved whole-genome metagenomic data



revealing bacterial domain predominance (97.4%–97.5%), with Proteobacteria (44.3%–50.0%) and Actinobacteria (22.1%–19.5%) as the dominant phyla—findings directly relevant to analogous Indian industrial soil contexts including DNH.

The WGS metagenomics workflow encompasses: (1) Shotgun library preparation by random DNA fragmentation (ultrasonication to 150–500 bp fragments), end-repair, A-tailing, adapter ligation, and size selection; (2) Paired-end sequencing on Illumina HiSeq/NovaSeq or long-read sequencing on Oxford Nanopore MinION/GridION; (3) Quality control using FastQC and Trimmomatic (adapter trimming, low-quality read removal, minimum length 100 bp); (4) Taxonomic profiling using Kraken2/Bracken (k-mer classification against RefSeq database), MetaPhlAn4 (marker gene-based), or DIAMOND against NCBI-NR; (5) Assembly using MEGAHIT or SPAdes for metagenomic contigs; (6) Functional annotation by Prokka (gene prediction), followed by mapping to KEGG, COG, and CARD databases using GhostKOALA, eggNOG-mapper, or DIAMOND; and (7) Statistical analysis in R using packages phyloseq, vegan, and DESeq2 (Bariya et al., 2025; Reddington et al., 2020; Vasudeva et al., 2022; Wei et al., 2021).

3.4 Functional Annotation Databases

3.4.1 KEGG (*Kyoto Encyclopedia of Genes and Genomes*)

The KEGG database maps metagenomic gene catalogues onto biochemical pathways, enabling reconstruction of community-level metabolic networks from environmental samples. In the Bariya *et al.* (2025) Silvassa study, KEGG annotation identified complete or near-complete metabolic pathways for benzoate degradation, chlorobenzoate catabolism, PAH metabolism, toluene and xylene degradation, chloroalkane and chloroalkene dehalogenation, and nitrotoluene and aminobenzoate catabolism. Vasudeva *et al.* (2022) similarly used GhostKOALA against the KEGG database to annotate 43.7% and 27.8% of open reading frames (ORFs) from activated sludge (BS14) and river water (SR1) samples respectively, revealing xenobiotic degradation modules with complete pathways alongside beta-lactam resistance genes.

3.4.2 COG (*Clusters of Orthologous Groups*)

COG-based annotation provides orthology-level functional classification of metagenomic proteins into 26 functional categories, enabling broad comparative assessment of community metabolic capabilities. In contaminated soil metagenomes from DNH and analogous Indian industrial zones, enrichment in COG category P (inorganic ion transport and metabolism, relevant to metal resistance), category C (energy production and conversion, reflecting metabolic adaptation to pollutant stress), and category Q (secondary metabolite biosynthesis, transport, and catabolism, encompassing many xenobiotic-degrading enzymes) is consistently reported (Bariya et al., 2025; Nipun et al., 2024; Zhang et al., 2022). The Feng et al. (2018) cadmium-contaminated soil study demonstrated that COG metabolic categories were the predominant functional annotations, with relative abundance of metabolic genes increased in Cd-contaminated compared to non-contaminated soils.

3.4.3 CARD (*Comprehensive Antibiotic Resistance Database*)

CARD analysis of metagenomic data enables systematic quantification of the soil resistome—the complete repertoire of antibiotic resistance genes. This is of particular relevance for industrial effluent-contaminated soils in DNH, where pharmaceutical manufacturing units contribute directly to selection pressure for ARG-carrying organisms, and where heavy metal contamination co-selects ARGs through shared mobile genetic elements. (de Abreu et al., 2021; Guiton et al., 2019; Vasudeva et al., 2022) documented 33 and 32 unique ARG types in activated sludge and river samples respectively using CARD, identifying beta-lactamase genes, efflux pump operons, and ribosomal protection proteins as predominant resistance mechanisms. Application of CARD to DNH soil metagenomes is essential for assessing the public health risk dimension of industrial soil contamination.

4. MICROBIAL COMMUNITY STRUCTURE IN INDUSTRIAL EFFLUENT-CONTAMINATED SOILS OF DNH

4.1 Taxonomic Profile: Phylum-Level Community Composition

Metagenomic investigations of industrial effluent-contaminated soils in Silvassa (DNH) reveal

a consistent pattern of community organization at the phylum level. Nipun *et al.* (2024), in the landmark first dedicated metagenomic study of DNH contaminated soils, reported dominance by Pseudomonadota (formerly Proteobacteria), Actinomycetota (formerly Actinobacteria), Bacillota (formerly Firmicutes), Myxococcota, Acidobacteriota, and Thermodesulfobacteriota across soil samples from Rakholi, Surangi, and Kherdi industrial estates. This phylum-level composition is congruent with findings from functionally analogous contaminated industrial soils across India: Sandhu *et al.* (2022) documented Proteobacteria (44.3%–50.0%) and Actinobacteria (22.1%–19.5%) as dominant phyla in PAH/PCB-contaminated soil near the Bhilai steel plant; Vasudeva *et al.* (2022) reported bacteria constituting 96% of the community in activated sludge from the Baddi CETP; and Tamil Nadu urban contaminated soils showed Proteobacteria enrichment under heavy metal and hydrocarbon co-contamination (Vasanthrao *et al.*, 2025).

The prevalence of Pseudomonadota as the dominant phylum across all DNH contaminated soil metagenomes reflects the exceptional metabolic versatility and stress tolerance of this phylogenetically diverse group. Pseudomonadota encompass aerobic and anaerobic metabolic strategies, fix atmospheric nitrogen, degrade a vast range of organic compounds, and encode diverse metal efflux and detoxification systems. Their enrichment in contaminated soils is further attributable to their copiotrophic lifestyle—efficient exploitation of the elevated organic carbon pools contributed by industrial effluents. Actinomycetota—the second most abundant phylum—contribute extensively to the degradation of recalcitrant organic compounds including PAHs, lignin derivatives, and chlorinated compounds through their diverse suite of oxidative enzymes (laccases, peroxidases, dioxygenases) (Das *et al.*, 2021; C. Su *et al.*, 2023; H. Su *et al.*, 2022).

The detection of Thermodesulfobacteriota (sulphate-reducing bacteria, SRB) and Myxococcota in DNH contaminated soils is particularly significant. SRB play critical roles in anaerobic heavy metal immobilization through biogenic sulphide production that precipitates metal sulphide

minerals, representing a natural bio-attenuation mechanism operating in anoxic micro-niches within contaminated soil aggregates. Myxococcota are predatory bacteria that regulate microbial community structure through bacterivory and contribute secondary metabolite biosynthesis capacity to contaminated soil ecosystems (Nipun *et al.*, 2024). Together, the phylum-level community architecture of DNH contaminated soils reflects a community assembled by multi-pollutant selection pressure, with different phyla occupying distinct ecological niches defined by oxygen availability, contaminant type, and physicochemical microenvironment.

4.2 Key Genera with Bioremediation Relevance

At the genus level, Bariya *et al.* (2025) identified *Pseudomonas* and *Bacillus* as the dominant genera across all three Silvassa contaminated soil samples (SA4, SA5, SA6), alongside *Sulfuricurvum kujiense*, *Novosphingobium sp.*, *Variovorax paradoxus*, *Usitatibacter rugosus*, and *Cupriavidus sp.* Nipun *et al.* (2024) additionally reported *Clostridium*, *Candidatus Entotheonella*, and *Desulforomonas* as abundant community members. Each of these genera possesses well-characterized bioremediation-relevant traits:

***Pseudomonas*:** This genus represents perhaps the most versatile xenobiotic-degrading bacterial genus known to science, encoding complete catabolic pathways for BTEX (benzene, toluene, ethylbenzene, xylene), naphthalene, phenanthrene, anthracene, polychlorinated biphenyls (PCBs), and diverse aliphatic hydrocarbons. *Pseudomonas aeruginosa* and *P. putida* produce biosurfactants (rhamnolipids) that increase the bioavailability of hydrophobic pollutants. Metal resistance is conferred by chromosomally and plasmid-encoded efflux systems (*czc*, *cop*, *mer* operons) (Banerjee *et al.*, 2022; Bariya *et al.*, 2025; Eze, 2021; Lee *et al.*, 2019; Mishra *et al.*, 2025).

***Bacillus*:** Spore-forming Firmicutes with exceptional desiccation, heat, and chemical tolerance, enabling persistence during seasonal dry periods in DNH's monsoon-dominated climate. *Bacillus* species produce biosurfactants (iturin, surfactin, fengycin), heteropolysaccharides that facilitate heavy metal biosorption, and extracellular enzymes (lipases, proteases, amylases) that



mineralize complex organic substrates. They also carry plasmid-borne arsenic, cadmium, and mercury resistance genes (Bariya et al., 2025; Nipun et al., 2024).

Cupriavidus sp.: The type species *Cupriavidus metallidurans* is the premier model organism for heavy metal resistance, encoding the chromosomal *czc* (cadmium-zinc-cobalt) resistance cluster, the plasmid-borne *chr* (chromate) operon, and *mer* (mercury) resistance. Its detection in DNH soils indicates adaptation to electroplating and tanning industry effluents. *Cupriavidus pinatubonensis* additionally degrades 2,4,5-trichlorophenoxyacetic acid (2,4,5-T) and other chlorinated herbicides (Bariya et al., 2025).

Novosphingobium sp.: Members of the Sphingomonadaceae family encode unique metabolic capabilities including degradation of lignin-derived aromatic monomers, PAHs, herbicides (atrazine, s-triazines), and diverse xenobiotic compounds via the modified ortho-cleavage pathway. Their characteristic sphingolipid cell membranes confer tolerance to organic solvent stress in effluent-contaminated soils (Bariya et al., 2025).

Variovorax paradoxus: An exceptionally metabolically versatile betaproteobacterium capable of lithoautotrophic growth on H₂, utilization of acyl-homoserine lactone quorum sensing molecules as sole carbon and energy source (limiting biofilm formation by pathogens), degradation of diverse xenobiotics including s-triazine herbicides and phenoxyacetic acid herbicides, and provision of plant growth-promoting traits in contaminated soils (Bariya et al., 2025; Ghimire et al., 2022; Huang et al., 2018).

Sulfuricurvum kujiense: A unique obligately anaerobic, autotrophic, sulphur-oxidizing epsilon-proteobacterium originally isolated from groundwater at a fuel-contaminated site. Its presence in DNH contaminated soils indicates the existence of anoxic micro-niches where sulphur redox cycling is active, with implications for the biological cycling of sulphur-containing industrial effluent compounds and heavy metal speciation through biogenic sulphide production (Bariya et al., 2025).

4.3 Effects of Industrial Effluent Contamination on Microbial Diversity

www.thebioscan.com

Industrial effluent contamination exerts measurable and well-documented effects on soil microbial diversity metrics. Alpha diversity—the richness and evenness of taxa within individual soil samples—consistently decreases with increasing contamination intensity in DNH-type industrial soil environments. Elevated concentrations of toxic heavy metals (Cr⁶⁺, Cd, Pb) suppress sensitive taxa through protein denaturation (binding of Cr⁶⁺ to sulfhydryl groups), disruption of DNA replication (Pb²⁺ competition with essential divalent cations), and reactive oxygen species generation (Cd²⁺ Fenton chemistry), effectively narrowing the diversity of viable community members to those encoding resistance mechanisms (Feng et al., 2018; Mi et al., 2025; C. Su et al., 2023).

Beta diversity—the difference in community composition between sites—provides a sensitive indicator of contamination-driven community restructuring. PCoA and NMDS ordination of 16S rRNA amplicon datasets from DNH industrial estates and adjacent reference soils are expected to reveal clear separation of contaminated and non-contaminated communities, consistent with findings from analogous Indian industrial soil investigations (Sandhu et al., 2022). Pollution load index (PLI), calculated from individual metal contamination factors, shows positive correlation with community dissimilarity from reference conditions—a relationship that can be leveraged to develop microbial community-based contamination bio-indices applicable to routine regulatory monitoring of DNH industrial zones.

Of particular ecological concern is the demonstrated suppression of functionally important microbial guilds in heavily contaminated soils. Arbuscular mycorrhizal fungi (AMF), nitrogen-fixing Rhizobia, and nitrifying *Nitrosomonas/Nitrobacter* communities—whose activities are essential for soil fertility and plant nutrition—show disproportionate sensitivity to heavy metal contamination, potentially contributing to the degradation of agricultural productivity in land areas adjacent to DNH's industrial estates (Mi et al., 2025; C. Su et al., 2023; H. Su et al., 2022).

4.4 Archaeal, Fungal, and Viral Community Components

While bacterial communities have dominated metagenomic investigations of DNH contaminated soils, archaea, fungi, and viruses constitute ecologically important components of these ecosystems. WGS metagenomics data from the Bhilai steel plant study (Sandhu et al., 2022) documented archaea at 1.2%–0.9% and eukaryotes at 1.4%–1.5% of total community reads, with viral sequences at 0.02%–0.04%, indicating that multi-kingdom community characterization is feasible from standard metagenomic datasets. In DNH contaminated soils, archaeal contributions to nitrogen cycling (ammonia-oxidizing archaea, AOA, of the Thaumarchaeota phylum) and sulphur metabolism (Crenarchaeota) are relevant to understanding contamination-driven disruption of biogeochemical cycles.

The fungal mycobiome of industrial effluent-contaminated soils harbors taxa with significant PAH biosorption and degradation capacity (*Aspergillus*, *Penicillium*, *Trichoderma*, *Rhizopus*) alongside white-rot Basidiomycetes (*Phanerochaete chrysosporium*, *Trametes versicolor*) whose ligninolytic peroxidase and laccase enzyme systems mediate non-specific oxidative attack on recalcitrant PAH structures. Pathak et al. (2020) demonstrated genus-level fungal community structuring by specific metal contaminants at mercury-polluted sites (*Penicillium* spp. positively correlated with total Hg; *Trichoderma* and *Aspergillus* correlated with methylmercury), a finding that motivates systematic mycobiome characterization of DNH's diverse heavy metal contamination profiles from electroplating (Cr, Ni, Cd) and tanning (Cr) industries.

The soil virome—bacteriophage and archaeal virus communities—influences microbial community structure through predation-driven top-down control (lysis of dominant bacterial taxa, enabling diversity maintenance through 'kill-the-winner' dynamics) and horizontal gene transfer of functional genes via transduction. In contaminated soils, phage-mediated transfer of ARGs and metal resistance genes between bacterial populations has been documented, with implications for the evolutionary trajectory of resistance gene dissemination in DNH's chronically contaminated soil ecosystems. Integration of viral metagenomic approaches (viromics) into future DNH soil

investigations is recommended to complete the community ecological picture (Nwachukwu & Babalola, 2022).

5. FUNCTIONAL METAGENOMICS: METABOLIC POTENTIAL OF DNH INDUSTRIAL SOIL MICROBIOME

5.1 Hydrocarbon Degradation Pathways

The most detailed functional metagenomic characterization of industrial effluent-contaminated soils in DNH to date was provided by Bariya et al. (2025), who identified a comprehensive aromatic hydrocarbon degradation metabolic network through KEGG annotation. The identified pathways collectively constitute a community-level 'degradome' capable of mineralizing structurally diverse organic pollutants from multiple industrial sources. The network architecture encompasses:

Benzoate and chlorobenzoate degradation: Benzoate is a central metabolic hub for aromatic compound catabolism. The ortho-cleavage pathway (catechol → cis,cis-muconate → β-ketoadipate → succinyl-CoA + acetyl-CoA) and meta-cleavage pathway (catechol → 2-hydroxymuconate semialdehyde → pyruvate + acetaldehyde) are both represented in DNH soil metagenomes, reflecting the metabolic diversity of *Pseudomonas*, *Acinetobacter*, and *Comamonas* populations that encode complementary pathway variants (Bariya et al., 2025).

Polycyclic aromatic hydrocarbon (PAH) catabolism: PAHs from plastic manufacturing and petroleum contamination are initiated by ring-hydroxylating dioxygenases (RHDs) that introduce two hydroxyl groups across aromatic double bonds, generating dihydrodiol intermediates. Naphthalene dioxygenase (NDO), biphenyl dioxygenase (BPDO), and phenanthrene dioxygenase systems identified in DNH soil metagenomes (Bariya et al., 2025) enable catabolism of 2–4 ring PAHs. Higher molecular weight PAHs (5–6 rings: benzo[a]pyrene, benzo[a]anthracene, coronene) require additional monooxygenase-mediated initial attack and represent a more recalcitrant fraction that may require fungal co-metabolism or extended incubation periods for significant mineralization (Lopez et al., 2022; Rahmeh et al., 2021).

BTEX (benzene, toluene, ethylbenzene, xylene) degradation: Toluene and xylene are



metabolized via toluene-4-monooxygenase (T4MO), toluene-2-monooxygenase, and benzyl succinate synthase (BSS) pathways under aerobic and anaerobic conditions respectively. Ethylbenzene catabolism proceeds via the *ebdABC* ethylbenzene dehydrogenase complex under anaerobic conditions, as relevant to anoxic micro-niches in DNH's waterlogged monsoon-season soils (Bariya et al., 2025; Hernández-Ospina et al., 2024).

Chlorinated compound dehalogenation: Haloalkane dehalogenase (*dhlA/linB*) and haloacid dehalogenase (HAD superfamily) genes detected in DNH soil metagenomes catalyse hydrolytic dehalogenation of chlorinated aliphatic—important contaminants from DNH's textile dyeing (chlorinated solvents) and pesticide-contaminated agricultural soils adjacent to industrial zones. Reductive dehalogenases from organohalide-respiring *Desulfitobacterium* and Dehalococcoides populations enable anaerobic transformation of polychlorinated compounds (Bariya et al., 2025; Muccee & Ejaz, 2020; Nipun et al., 2024).

5.2 Heavy Metal Resistance Gene Repertoire

Functional annotation of DNH contaminated soil metagenomes by Bariya *et al.* (2025) and Nipun *et al.* (2024) reveals a diverse arsenal of heavy metal resistance genes whose distribution reflects the specific industrial effluent contamination profile of different sites. Key metal resistance systems identified in DNH and analogous Indian contaminated soil metagenomes include: (i) *czc* (cadmium-zinc-cobalt resistance): *czcA* efflux pump and *czcD* facilitator encoded in *Cupriavidus*, *Ralstonia*, and *Comamonas* populations; (ii) *chr* (chromate resistance): *chrA* chromate transporter that actively exports Cr^{6+} from the cytoplasm, relevant to electroplating industry-contaminated sites in DNH; (iii) *ars* (arsenic resistance): *arsRBC* operon (arsenate reductase *arsC*, arsenite exporter *arsB*, transcriptional repressor *arsR*) and AIO arsenite oxidase, relevant where arsenic from agricultural pesticide use co-contaminates industrial zones; (iv) *mer* (mercury resistance): *merA* mercuric ion reductase and *merB* organomercury lyase, relevant to pharmaceutical industry sites in DNH; and (v) *cop* (copper resistance): *copA/copB* P-type ATPase efflux pumps, relevant to electroplating

industry sites (Bera et al., 2025; Das et al., 2021; Feng et al., 2018; Mi et al., 2025).

Redundancy analysis (RDA) of metal resistance gene abundance versus soil metal concentrations in analogous contaminated soil studies consistently identifies strong positive correlations between specific metal concentrations and their corresponding resistance gene abundances, confirming metal-driven selection of resistance determinants. In DNH's multi-industry contamination landscape, RDA is expected to reveal distinct resistance gene profiles across different industrial estate types—chromate/nickel resistance gene enrichment near electroplating units, cadmium/lead resistance near battery recycling operations, and copper resistance near electrical manufacturing facilities (Cruz-Hernández et al., 2022; Das et al., 2021; Mi et al., 2025).

5.3 Antibiotic Resistance Gene (ARG) Co-Selection in Industrial Soil

A critically important finding emerging from metagenomic analysis of industrial effluent-contaminated soils is the co-selection of ARGs alongside metal resistance genes through shared mobile genetic elements. Vasudeva *et al.* (2022) documented 33 unique ARG types in activated sludge from the Baddi CETP using CARD analysis, including beta-lactamase genes (*blaTEM*, *blaSHV*), efflux pump operons (*mexAB-oprM*, *acrAB-tolC*), and aminoglycoside modifying enzymes. This ARG reservoir, maintained by industrial effluent selection pressure in the absence of direct antibiotic application, constitutes a public health threat through the environmental dissemination of resistance determinants that could transfer to pathogenic bacteria (Vasanthrao et al., 2025).

Two complementary mechanisms drive ARG-metal resistance co-selection in DNH industrial soils: co-resistance (ARGs and MRGs physically co-located on the same plasmid, integron, or transposon) and cross-resistance (a single resistance mechanism—typically an efflux pump or altered membrane permeability—confers tolerance to both heavy metals and structurally unrelated antibiotics). DNH's pharmaceutical manufacturing sector introduces a third dimension: direct pharmaceutical antibiotic contamination of soils through effluent discharge, creating combined



antibiotic-metal selection pressure that may accelerate the evolution and consolidation of multi-drug-resistant, metal-tolerant bacterial lineages in industrial soil ecosystems. Metagenomic resistome profiling using CARD and MobileElementFinder tools is urgently warranted for comprehensive risk assessment across DNH's diverse industrial estate types.

5.4 Nitrogen and Sulphur Cycling Genes in Contaminated Soils

Beyond xenobiotic degradation and metal resistance, functional metagenomics reveals contamination-driven perturbations of essential biogeochemical cycling pathways in DNH industrial soils. Nitrogen cycling—encompassing nitrogen fixation (*nifH*), ammonia oxidation (*amoA*), nitrite oxidation (*nxrAB*), denitrification (*nirS*, *nirK*, *nosZ*), and anaerobic ammonia oxidation (*hzsABC*)—is particularly sensitive to heavy metal inhibition, as metal ions competitively inhibit the metal-containing active sites of key enzymes (Ni in urease, Mo/Fe in nitrogenase, Cu in multicopper oxidases). Disruption of nitrogen cycling in contaminated DNH soils may cascade into agricultural nutrient deficits in land areas irrigated with contaminated Daman Ganga water (Mi et al., 2025; C. Su et al., 2023).

Sulphur cycling genes, including sulphate reduction (*dsrAB*), sulphur oxidation (*soxABCDXYZ*), and sulphur disproportionation enzymes, are enriched in DNH contaminated soils that harbor *Thermodesulfobacteriota* and *Sulfuricurvum* populations (Nipun et al., 2024). The coupled sulphur–metal cycle is of direct relevance to in situ bioremediation: biogenic sulphide produced by sulphate-reducing bacteria precipitates metal ions as insoluble metal sulphide minerals (CdS, PbS, ZnS), effectively immobilizing labile metal fractions and reducing their bioavailability and plant/invertebrate toxicity. Functional metagenomic quantification of *dsrAB* gene abundance can serve as a proxy for the biotic sulphide production capacity of contaminated soils, informing decisions about whether sulphate amendment-based bioremediation strategies would be effective in specific DNH sites.

6. FACTORS SHAPING MICROBIAL COMMUNITY STRUCTURE IN DNH INDUSTRIAL SOILS

6.1 Soil Physicochemical Parameters as Community Drivers

Redundancy analysis (RDA) and Mantel tests from contaminated soil metagenomics studies consistently identify soil pH, total organic carbon (TOC), electrical conductivity (EC), and individual heavy metal concentrations as the primary environmental variables explaining between-site variation in microbial community composition (Mi et al., 2025; Pérez-De-Mora et al., 2006; H. Su et al., 2022). In DNH industrial soils, pH modulation by acidic effluents from electroplating (HCl, H₂SO₄ pickling baths) and textile dyeing units creates spatially heterogeneous pH gradients that serve as primary filters on bacterial community membership. Soil pH declines from neutral to acidic conditions (pH 4.5–6.0) suppress sensitive neutrophilic Bacteroidetes and shift community composition toward acid-tolerant *Acidobacteria* and Alphaproteobacteria lineages.

Elevated TOC in soils receiving high-BOD organic effluents from food processing, textile, and chemical units supports copiotrophic Proteobacteria populations while potentially immobilizing heavy metals through organic-metal complexation, reducing their immediate bioavailability and toxic impact. However, organic-metal complexes can be remobilized upon mineralization of the organic fraction, creating episodic metal toxicity pulses that challenge community stability. EC elevation from inorganic salt loading in industrial effluents (from textile auxiliaries, electroplating baths, cooling water blowdown) suppresses salt-sensitive taxa and selects for halotolerant lineages, further reducing community diversity in areas receiving concentrated effluent discharges (Bariya et al., 2025; Feng et al., 2018; Pandit et al., 2020).

6.2 Contamination Type and Source-Specific Community Structuring

A defining characteristic of the DNH industrial landscape is the co-occurrence of chemically diverse industrial effluent types within small geographic areas, creating overlapping contamination footprints with complex synergistic and antagonistic contaminant interactions. Metagenomically distinct community profiles are expected across sites influenced by different industrial effluent types: electroplating sites (Cr, Ni,



Cd contamination) select for metal resistance gene-enriched communities with *Cupriavidus*, *Ralstonia*, and *Lysinibacillus* as dominant metal-tolerant genera; textile dyeing sites (organic dye, BOD, TOC contamination) enrich dye-decolorising *Pseudomonas*, *Bacillus*, and fungal populations; pharmaceutical sites (API, solvent contamination) select for cytochrome P450-encoding strains (*Streptomyces*, *Mycobacterium*) capable of oxidative biotransformation of pharmaceutical compounds (Patel et al., 2016; Sandhu et al., 2022; Vasudeva et al., 2022).

Spatial metagenomics—systematic collection of samples along transects from contamination sources (effluent discharge points, waste storage areas) to reference zones—would enable construction of contamination-response models for DNH's industrial effluent microbial ecology. Such models could identify threshold contamination levels above which community diversity collapses and bioremediation potential is compromised—essential information for guiding regulatory enforcement priorities and remediation investment targeting at the most ecologically damaged sites (Nipun et al., 2024; C. Su et al., 2023).

6.3 Seasonal Dynamics in the Monsoon-Dominated Climate

DNH's climate is dominated by an intense southwest monsoon (June–September) delivering 2,000–2,500 mm annual rainfall, followed by a dry winter (October–February) and a hot, dry pre-monsoon season (March–May). This pronounced seasonal precipitation regime drives dramatic oscillations in soil moisture, temperature, redox potential, and contaminant mobility that impose strong temporal structuring on microbial communities. During the monsoon, soil waterlogging creates anaerobic micro-niches that support sulphate-reducing bacteria (Thermodesulfobacteriota, *Desulforomonas*), methanogens, and organohalide-respiring populations. Contaminant mobilization by rainwater percolation redistributes heavy metals and organic pollutants within and between industrial estates, exposing previously uncontaminated soils to pollution and altering the spatial pattern of microbial community disruption.

www.thebioscan.com

Pre-monsoon heat and desiccation select for sporulating *Bacillus* and *Clostridium* populations (detected in DNH soils by (Bariya et al., 2025; Nipun et al., 2024), and desiccation-tolerant Actinomycetota lineages. Post-monsoon rewetting triggers community resuscitation events where dormant populations resume metabolic activity simultaneously, potentially driving rapid community turnover. Longitudinal 16S rRNA amplicon monitoring across multiple seasonal cycles would be essential to disentangle contamination-driven from climate-driven community dynamics in DNH soils—a critical precondition for reliable interpretation of any cross-sectional metagenomic survey data and for designing bioremediation interventions that account for seasonally variable community composition and metabolic activity.

6.4 Industrial Effluent Versus Agricultural Soil Context

Many DNH industrial zones are geographically proximate to or intermixed with agricultural lands where local Adivasi (indigenous tribal) communities practice subsistence and commercial farming. Contaminated river water from the Daman Ganga and its tributaries is used for irrigation, introducing industrial pollutants into agricultural soils and potentially disrupting the rhizosphere microbial communities essential for crop plant nutrition and health. Patel *et al.* (2016) demonstrated resilience and adaptation responses in soil microbial communities exposed to edible oil contamination from industrial estates—a scenario directly relevant to agricultural soils adjacent to food processing units in DNH. Understanding the continuum between industrial soil contamination and agricultural soil impact through metagenomic monitoring of both soil types would provide a complete picture of the ecological footprint of DNH's industrial effluent discharges.

7. METAGENOMICS-GUIDED BIOREMEDIATION STRATEGIES FOR DNH INDUSTRIAL SOILS

7.1 Bio-stimulation of Indigenous Degradable Communities

Bio-stimulation leverages the metabolically competent indigenous microbiome—whose richness has been partially characterized by Nipun et al. (2024) and Bariya et al. (2025)—by supplying rate-



limiting nutrients, electron donors, or electron acceptors that overcome the nutritional and geochemical constraints on in situ pollutant degradation. For petroleum hydrocarbon-contaminated DNH soils, nitrogen and phosphorus supplementation to achieve a C:N:P ratio of 100:10:1 is the primary bio-stimulation strategy, addressing the nutrient limitations imposed by hydrocarbon-rich but nitrogen-poor effluent-affected soils. Oxygen delivery through bioventing or biosparging can overcome aerobic respiration limitations in deeper, saturated zones.

Metagenomics provides essential guidance for bio-stimulation by identifying which nutrient limitations are most constraining for degrader populations. If functional annotation reveals abundance of nitrogen fixation genes (*nifH*) relative to nitrification genes (*amoA*), this indicates that diazotrophic nitrogen fixers are already providing some nitrogen input, and that phosphorus limitation may be more critical than nitrogen limitation for supplementation. Similarly, quantification of *dsrAB* sulphate reduction genes versus RHD aromatic degradation genes can indicate whether sulphur-cycling bio-stimulation (sulphate addition to stimulate SRB-mediated metal immobilization) or organic carbon supplement (to stimulate aerobic hydrocarbon degradation) would yield greater remediation benefits at specific DNH sites (Yadav et al., 2022)

7.2 Bioaugmentation with Indigenous Strain Consortia

Bioaugmentation—the addition of exogenous microbial cultures to enhance pollutant degradation—is most effective when the introduced strains are sourced from the same contaminated ecosystem, ensuring pre-adaptation to the physicochemical conditions and contaminant profile of the target site. The metagenomic inventories of DNH contaminated soils generated by Nipun *et al.* (2024) and Bariya *et al.* (2025) provide the genus-level roadmap for targeted isolation of bioaugmentation candidates: *Pseudomonas*, *Bacillus*, *Cupriavidus*, *Novosphingobium*, and *Variovorax* can be selectively cultivated on mineral salt medium with site-specific pollutants as sole carbon sources or electron donors, then characterized by whole-genome sequencing to confirm the

presence of relevant catabolic and resistance gene clusters before deployment.

Consortium bioaugmentation—inoculation of multiple complementary strains with non-overlapping catabolic capabilities—is generally more effective than single-strain approaches for the complex multi-pollutant environments of DNH industrial soils. A hypothetical DNH bioaugmentation consortium might include: an aerobic *Pseudomonas* strain for BTEX and PAH degradation; a *Bacillus* strain for biosurfactant production and bioavailability enhancement; a *Cupriavidus* strain for chromate and cadmium resistance and sequestration; and a *Desulforomonas* or *Desulfovibrio* strain for anaerobic metal sulphide precipitation in waterlogged zones. Metatranscriptomic monitoring post-bioaugmentation would confirm in situ expression of the introduced strains' target catabolic genes, providing mechanistic evidence of bioaugmentation effectiveness beyond bulk pollutant concentration measurements (Eze, 2021; Hernández-Ospina et al., 2024; Kumar et al., 2024).

7.3 Phytoremediation-Microbiome Synergy

Phytoremediation—plant-mediated contaminant extraction, degradation, or stabilization—operates synergistically with root-associated microbial communities that are continuously stimulated by rhizosphere carbon input from root exudates (sugars, organic acids, amino acids, root-border cells). The rhizosphere effect—a 10–100-fold enrichment of microbial biomass and metabolic activity relative to bulk soil—can be specifically channeled toward pollutant-degrading community members by selecting plant species that exude compounds (salicylate, coumarin, flavonoids) that induce the expression of catabolic genes in associated bacteria. Metagenomics enables systematic characterization of rhizosphere microbiomes of plant species naturally colonizing DNH contaminated soils, identifying the most promising plant-microbe combinations for phytoremediation deployment.

Sun et al. (2023) demonstrated that metagenomics-guided construction of composite microbial agents could accelerate crude oil biodegradation in contaminated soil systems by reconstructing a functionally complementary



microbiome whose members engage in syntrophic electron transfer and metabolic cross-feeding. Applying this approach to DNH contaminated soils would involve: (1) metagenomic identification of the keystone degrader genera from Nipun *et al.* (2024) and Bariya *et al.* (2025) datasets; (2) selective cultivation and characterization of these genera; (3) formulation of a composite bioinoculant with a plant growth-promoting rhizobacterium (PGPR) carrier strain; and (4) greenhouse and field trials measuring both pollutant removal rates and plant biomass production as co-benefits of the integrated phytoremediation-bioaugmentation strategy (Singh *et al.*, 2022).

7.4 Metagenome-Assembled Genomes (MAGs) and Novel Enzyme Discovery

Assembly-based metagenomic analysis—reconstructing metagenome-assembled genomes (MAGs) from co-assembled metagenomic reads—represents the frontier of functional discovery in contaminated soil ecosystems. MAG reconstruction enables: (i) assignment of specific functional genes (including novel, uncharacterized catabolic genes) to defined genomic contexts within specific taxonomic lineages; (ii) analysis of metabolic pathway completeness within individual genome bins; (iii) identification of previously uncultured lineages with novel combinations of metal resistance and xenobiotic degradation capabilities; and (iv) discovery of novel enzyme families through sequence-function correlation approaches applied to environmental genomic data (Nwachukwu & Babalola, 2022; Sousa *et al.*, 2022).

The diversity and persistence of industrial pollutant selection pressures in DNH soils over several decades of industrialization have likely driven significant evolutionary innovation in catabolic enzyme families—potential sources of industrial biotechnology applications. Deep shotgun sequencing (≥ 20 Gbp per sample) of DNH contaminated soil samples, followed by binning with DAS_Tool or MetaBAT2 and quality assessment with CheckM (completeness $\geq 70\%$, contamination $\leq 10\%$), would generate a draft MAG catalogue of DNH's industrial soil microbiome. Functional annotation of novel open reading frames in these MAGs, followed by heterologous expression in *E. coli* or *Pseudomonas* chassis strains, could unlock

www.thebioscan.com

DNH-derived enzymes with industrial applications in green chemistry, sustainable textiles, and environmental engineering.

8. CHALLENGES, RESEARCH GAPS, AND FUTURE PERSPECTIVES

8.1 Technical and Methodological Challenges

Several significant technical challenges constrain the rigor and reproducibility of metagenomic investigations in DNH industrial soils. Sample representativeness is paramount but difficult to achieve given the high spatial heterogeneity of contamination patterns across industrial estates of varying effluent composition, operational history, and proximity to drainage infrastructure. Composite sampling strategies (pooling 5–10 sub-samples per plot, collected at multiple depths: 0–10 cm, 10–30 cm, 30–50 cm) combined with adequate biological replication (minimum $n=3$ per site and contamination category) and statistically powered sampling design are essential for generating community composition data that reflects site-level rather than point-source variability.

The co-extraction of PCR inhibitors from chemically complex industrial soils—humic acids, metal ions, industrial solvents—remains a persistent practical challenge despite advances in extraction kit chemistry. Validation of DNA extraction efficiency using internal spike-in controls (defined quantities of *Thermus thermophilus* DNA added before extraction, quantified after extraction by qPCR) is recommended for all DNH soil metagenomic studies to enable cross-sample yield normalization and bias assessment. Bioinformatic challenges include the low percentage of reads that can be assigned to characterized reference sequences when analyzing environments with high proportions of novel, uncharacterized taxa—estimated at 20–40% unknown sequences in industrial contaminated soil metagenomes—necessitating complementary de novo assembly and MAG-based approaches (Bariya *et al.*, 2025; Li *et al.*, 2023; Xu *et al.*, 2022).

8.2 Multi-Omics Integration for Mechanistic Understanding

Metagenomics characterizes community genomic potential—the complete gene repertoire—but cannot directly measure which genes are actually expressed, at what levels, and under what conditions. Integration of meta transcriptomics (sequencing of



community mRNA to quantify gene expression), meta proteomics (mass spectrometry-based identification and quantification of expressed proteins), and metabolomics (targeted and untargeted profiling of small molecule metabolites) with metagenomics provides a multi-dimensional portrait of community function that is essential for translating genomic discoveries into operational bioremediation interventions (Dhar et al., 2023; Sousa et al., 2022).

For DNH industrial soils, a proposed multi-omics investigation framework would encompass: (1) simultaneous co-extraction of DNA and RNA (using PowerSoil DNA Elution Accessory Kit for RNA; DNase treatment of RNA fraction) for paired metagenomics and Meta transcriptomics from the same samples; (2) metaproteomic extraction using bead-beating lysis and SDS-PAGE peptide fractionation, with LC-MS/MS identification against DNH soil metagenome-derived protein databases; (3) GC-MS/LC-MS metabolomics of soil extracts to quantify pollutant intermediates and end products; and (4) statistical co-inertia analysis (MFA, MOFA+) to identify cross-omics correlations between gene expression patterns and pollutant transformation rates (Gutierrez Reyes, C. D., Alejo-Jacuinde, G., Perez Sanchez, B., Chavez Reyes, J., Onigbinde, S., Mogut, D., ... & Mechref, 2024; Mishra et al., 2025; Vasudeva et al., 2022).

8.3 Long-Read Sequencing and Technological Advances

The adoption of third-generation long-read sequencing platforms—Oxford Nanopore Technology (ONT) MinION/PromethION and Pacific Biosciences (PacBio) SMRT sequencing—is increasingly transforming environmental metagenomics. Long reads (10–100 kb average length for ONT; 15–25 kb for PacBio HiFi) resolve repetitive genomic regions, complete circular plasmid sequences in a single read, enable direct detection of base modifications (methylation patterns, indicating epigenetic regulation of resistance genes), and dramatically improve MAG completeness and contamination scores compared to short-read assemblies. Sandhu et al. (2022) demonstrated the utility of Oxford Nanopore GridION for whole-genome metagenomic analysis of Indian industrial contaminated soils, providing a

www.thebioscan.com

precedent for adoption of this technology in future DNH investigations.

Hybrid sequencing approaches combining Illumina short reads (for high base accuracy) with ONT long reads (for scaffolding and repeat resolution) currently represent the gold standard for metagenomic assembly quality. The falling cost trajectory of ONT sequencing, combined with the availability of portable MinION devices that could be operated at field stations adjacent to DNH industrial sites, opens the prospect of near-real-time in situ metagenomic monitoring—a transformative capability for regulatory compliance monitoring and bioremediation progress tracking. MinION-based sequencing of 16S rRNA amplicons at industrial effluent discharge points in DNH could provide same-day taxonomic community data to guide regulatory enforcement decisions (Nwachukwu & Babalola, 2022; Reddington et al., 2020; Sandhu & Paul, 2021).

8.4 Policy Framework and Regulatory Integration

The scientific value of metagenomic investigations of DNH industrial soils extends beyond academic knowledge generation to inform evidence-based environmental policy and regulatory frameworks. India is currently developing a National Soil and Sediment Metagenomics Policy drawing from international frameworks (EU Horizon Europe Soil Mission, USDA One Health Framework, China's Soil Pollution Action Plan, FAO Global Soil Partnership), with emphasis on AI-enabled monitoring infrastructure, standardized protocols, and reformed intellectual property frameworks for environmental microbial resource management (Dey et al., 2026). DNH, with its high industrial density, available baseline metagenomic data (Bariya et al., 2025; Nipun et al., 2024), and co-location of research institutions (NAMO MERI Silvassa, Dr. APJ Abdul Kalam Government College Silvassa, Naran Lala College of Professional and Applied Science), is well-positioned to serve as a model implementation site for India's emerging soil microbiome monitoring policy.

Practical policy recommendations arising from this review include: (i) mandatory pre-industrial and post-discharge metagenomic soil monitoring as a condition of environmental



clearance for new industrial projects in DNH; (ii) development of microbial community-based Ecological Status Classification criteria (analogous to EU Water Framework Directive biological quality elements) for DNH industrial soils using Shannon diversity, Proteobacteria: Actinomycetota ratios, and functional gene richness as index components; (iii) establishment of a dedicated DNH Soil Microbiome Observatory with shared sequencing infrastructure accessible to local research institutions; and (iv) integration of ARG surveillance data from DNH soil metagenomes into India's National AMR Action Plan to quantify the environmental contribution of industrial contamination to the domestic AMR burden.

9. CONCLUSION

This review has comprehensively examined metagenomic approaches and their application to understanding the microbial ecology of industrial effluent-contaminated soils across Dadra and Nagar Haveli—a Union Territory at the intersection of rapid industrialization and significant ecological vulnerability. The synthesis of primary metagenomic investigations from Silvassa and contextual findings from analogous Indian industrial soil systems supports the following principal conclusions:

First, industrial effluent-contaminated soils of Dadra and Nagar Haveli harbor taxonomically diverse microbial communities structurally dominated by Pseudomonadota, Actinomycetota, Bacillota, Myxococcota, Acidobacteriota, and Thermodesulfobacteriota, with key genera including *Pseudomonas*, *Bacillus*, *Cupriavidus*, *Novosphingobium*, *Variovorax*, *Sulfuricurvum*, and *Clostridium* collectively encoding extraordinary catabolic versatility and metal resistance capacity adapted to the territory's specific industrial effluent contamination profile.

Second, whole-genome shotgun metagenomics with KEGG and COG functional annotation has revealed a comprehensive, community-encoded degradation network for benzoate, PAHs, BTEX compounds, chlorinated aliphatics, nitrotoluenes, and aminobenzoates in Silvassa contaminated soils, with pathways converging on central metabolic intermediates (TCA cycle). This genomic evidence demonstrates that the indigenous soil microbiome possesses significant,

though currently largely untapped, in situ bioremediation potential.

Third, CARD-based resistome analysis of industrial effluent-contaminated soils from comparable Indian sites reveals substantial ARG co-selection alongside metal resistance genes through plasmid- and transposon-borne mobile genetic elements—an emerging public health concern that demands systematic resistome characterization of DNH industrial soils as part of India's National AMR Action Plan.

Fourth, metagenomics-guided bioremediation strategies—including nutrient-targeted bio stimulation, indigenous strain consortium bioaugmentation, and phytoremediation-microbiome integration—represent scientifically grounded and ecologically appropriate remediation pathways for DNH industrial sites, with the metagenomically characterized indigenous biota serving as both the diagnostic framework and the biological resource for interventions.

Fifth, significant knowledge gaps remain: comprehensive spatial and seasonal metagenomics across DNH's diverse industrial estate types; multi-omics integration linking genomic potential with transcriptional activity and metabolite dynamics; mycobiome and virome characterization; MAG-based novel enzyme discovery; and longitudinal monitoring of bioremediation interventions. Addressing these gaps requires coordinated investment in sequencing infrastructure, bioinformatics capacity, and collaborative research networks centered on DNH's locally embedded research institutions.

The ecological and public health stakes of industrial effluent soil contamination in Dadra and Nagar Haveli are profound, affecting the health of local communities, agricultural livelihoods, Daman Ganga River ecosystem integrity, and groundwater quality. Metagenomics provides the scientific foundation for a data-driven transition from reactive, end-of-pipe contamination management to proactive, microbiome-informed soil ecosystem stewardship—a transition that is both scientifically urgent and developmentally essential for one of India's most rapidly industrializing Union Territories.

10. DECLARATIONS



Conflict of Interest: The authors declare no conflict of interest.

Funding: This research received National Fellowship Scheme for Higher Education of ST students from Ministry of Tribal Affairs, India.

Data Availability: Not applicable (review article; no new data generated).

11. REFERENCES

- Ahmad, T., Gupta, G., Sharma, A., Kaur, B., El-Sheikh, M. A., & Alyemeni, M. N. (2021). Metagenomic analysis exploring taxonomic and functional diversity of bacterial communities of a Himalayan urban fresh water lake. *PLOS ONE*, 16(3), e0248116. <https://doi.org/10.1371/journal.pone.0248116>
- Allamin, I. A., Halmi, M. I. E., Yasid, N. A., Ahmad, S. A., Abdullah, S. R. S., & Shukor, Y. (2020). Rhizodegradation of Petroleum Oily Sludge-contaminated Soil Using *Cajanus cajan* Increases the Diversity of Soil Microbial Community. *Scientific Reports*, 10(1), 1–11. <https://doi.org/10.1038/s41598-020-60668-1>
- Banerjee, S., Bedics, A., Harkai, P., Kriszt, B., Alpula, N., & Tancsics, A. (2022). Evaluating the aerobic xylene-degrading potential of the intrinsic microbial community of a legacy BTEX-contaminated aquifer by enrichment culturing coupled with multi-omics analysis: uncovering the role of *Hydrogenophaga* strains in xylene degradation. *Environmental Science and Pollution Research*, 29(19), 28431–28445. <https://doi.org/10.1007/s11356-021-18300-w>
- Bariya, N., Dholaria, M., & Tailor, A. J. (2025). Hydrocarbon Degradation Potential Revealed by Metagenomic Analysis of Contaminated Soil Samples Across Silvassa, India: A Potential Bioremediation Solution. *Journal of Pure and Applied Microbiology*, 19(3), 1834–1853. <https://doi.org/10.22207/JPAM.19.3.09>
- Bera, A., Patel, R. R., Precilla, S. D., Singh, M., Agiesh Kumar, B., Bhowmik, S., Varshney, M., & Roy, S. (2025). External stimuli responsive syneresis of amino acid-based bioactive hydrogels: a sustainable platform for environmental remediation. *Materials Advances*, 6(15), 5184–5195. <https://doi.org/10.1039/D5MA00274E>
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., ... Caporaso, J. G. (2018). *QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science*. <https://doi.org/10.7287/peerj.preprints.27295v2>
- Cruz-Hernández, M. A., Mendoza-Herrera, A., Bocanegra-García, V., & Rivera, G. (2022). *Azospirillum* spp. from Plant Growth-Promoting Bacteria to Their Use in Bioremediation. *Microorganisms*, 10(5), 1057. <https://doi.org/10.3390/microorganisms10051057>
- Das, N., Kotoky, R., Maurya, A. P., Bhuyan, B., & Pandey, P. (2021). Paradigm shift in antibiotic-resistome of petroleum hydrocarbon contaminated soil. *Science of the Total Environment*, 757(xxxx), 143777. <https://doi.org/10.1016/j.scitotenv.2020.143777>
- de Abreu, V. A. C., Perdigão, J., & Almeida, S. (2021). Metagenomic Approaches to Analyze Antimicrobial Resistance: An Overview. *Frontiers in Genetics*, 11. <https://doi.org/10.3389/fgene.2020.575592>
- Dey, P., Jadhav, R. S., Kulshrestha, S., Hata, A., & Mukherjee, S. (2026). An overview of soil and sediment metagenomics policy: insights from India and global scenario. *Proceedings of the Indian National Science Academy*. <https://doi.org/10.1007/s43538-026-00688-z>
- Dhar, R., Basu, S., Bhattacharyya, M., & Dutta, T. K. (2023). Evaluation of distinct molecular architectures and coordinated regulation of the catabolic pathways of oestrogenic dioctyl phthalate isomers in *Gordonia* sp. *Microbiology (United Kingdom)*, 169(6), 1–19. <https://doi.org/10.1099/mic.0.001353>
- Eze, M. O. (2021). Metagenome Analysis of a Hydrocarbon-Degrading Bacterial Consortium Reveals the Specific Roles of BTEX



- Biodegraders. *Genes*, 12(1), 98. <https://doi.org/10.3390/genes12010098>
- Feng, G., Xie, T., Wang, X., Bai, J., Tang, L., Zhao, H., Wei, W., Wang, M., & Zhao, Y. (2018). Metagenomic analysis of microbial community and function involved in cd-contaminated soil. *BMC Microbiology*, 18(1), 1–13. <https://doi.org/10.1186/s12866-018-1152-5>
- Ghimire, N., Kim, B., Lee, C.-M., & Oh, T.-J. (2022). Comparative genome analysis among *Variovorax* species and genome guided aromatic compound degradation analysis emphasizing 4-hydroxybenzoate degradation in *Variovorax* sp. PAMC26660. *BMC Genomics*, 23(1), 375. <https://doi.org/10.1186/s12864-022-08589-3>
- Guitor, A. K., Raphenya, A. R., Klunk, J., Kuch, M., Alcock, B., Surette, M. G., McArthur, A. G., Poinar, H. N., & Wright, G. D. (2019). Capturing the Resistome: a Targeted Capture Method To Reveal Antibiotic Resistance Determinants in Metagenomes. *Antimicrobial Agents and Chemotherapy*, 64(1). <https://doi.org/10.1128/AAC.01324-19>
- Gutierrez Reyes, C. D., Alejo-Jacuinde, G., Perez Sanchez, B., Chavez Reyes, J., Onigbinde, S., Mogut, D., ... & Mechref, Y. (2024). Multi omics applications in biological systems. *Current Issues in Molecular Biology*, 43(6), 5777–5793.
- Hernández-Ospina, D. A., Osorio-González, C. S., Miri, S., & Kaur Brar, S. (2024). New perspectives on the anaerobic degradation of BTEX: Mechanisms, pathways, and intermediates. *Chemosphere*, 361, 142490. <https://doi.org/10.1016/j.chemosphere.2024.142490>
- Huang, Y., Xiao, L., Li, F., Xiao, M., Lin, D., Long, X., & Wu, Z. (2018). Microbial Degradation of Pesticide Residues and an Emphasis on the Degradation of Cypermethrin and 3-phenoxy Benzoic Acid: A Review. *Molecules*, 23(9), 2313. <https://doi.org/10.3390/molecules23092313>
- Hugenholtz, P. (2002). *Exploring prokaryotic diversity in the genomic era*. 1–8.
- Indiastat. (2022). *Number of polluted river stretches identified by CPCB in Dadra and Nagar Haveli and Daman and Diu*. CPCB. <https://doi.org/https://www.indiastat.com/dadra-and-nagar-haveli-and-daman-and-diu-state/data/environment-and-pollution>
- Jiang, L., Luo, C., Zhang, D., Song, M., Sun, Y., & Zhang, G. (2018). Biphenyl-Metabolizing Microbial Community and a Functional Operon Revealed in E-Waste-Contaminated Soil. *Environmental Science and Technology*, 52(15), 8558–8567. <https://doi.org/10.1021/acs.est.7b06647>
- Kumar, V., Ameen, F., & Verma, P. (2024). Unraveling the shift in bacterial communities profile grown in sediments co-contaminated with chlorolignin waste of pulp-paper mill by metagenomics approach. *Frontiers in Microbiology*, 15. <https://doi.org/10.3389/fmicb.2024.1350164>
- Lee, Y., Lee, Y., & Jeon, C. O. (2019). Biodegradation of naphthalene, BTEX, and aliphatic hydrocarbons by *Paraburkholderia* aromaticivorans BN5 isolated from petroleum-contaminated soil. *Scientific Reports*, 9(1), 860. <https://doi.org/10.1038/s41598-018-36165-x>
- Li, S., Lian, W.-H., Han, J.-R., Ali, M., Lin, Z.-L., Liu, Y.-H., Li, L., Zhang, D.-Y., Jiang, X.-Z., Li, W.-J., & Dong, L. (2023). Capturing the microbial dark matter in desert soils using culturomics-based metagenomics and high-resolution analysis. *Npj Biofilms and Microbiomes*, 9(1), 67. <https://doi.org/10.1038/s41522-023-00439-8>
- Lopez, E. S., Elufisan, T. O., Bustos, P., Charles, C. P. M., Mendoza-Herrera, A., & Guo, X. (2022). Complete Genome Report of a Hydrocarbon-Degrading *Sphingobium yanoikuyae* S72. *Applied Sciences*, 12(12), 6201. <https://doi.org/10.3390/app12126201>
- Mi, Q., Wu, Y., Cai, H., Xu, Z., Zhao, Y., Guan, R., Fan, X., & Guo, J. (2025). Effects of Long-Term Heavy Metal Pollution on Microbial Community Structure in Soil. *Toxics*, 13(9), 806. <https://doi.org/10.3390/toxics13090806>
- Mishra, A. K., Yadav, P., Sharma, S., & Maurya, P. (2025). Comparison of microbial diversity and community structure in soils managed with organic and chemical fertilization strategies using amplicon sequencing of 16 s and ITS



- regions. *Frontiers in Microbiology*, 15. <https://doi.org/10.3389/fmicb.2024.1444903>
- Mucce, F., & Ejaz, S. (2020). Whole genome shotgun sequencing of POPs degrading bacterial community dwelling tannery effluents and petrol contaminated soil. *Microbiological Research*, 238(May), 126504. <https://doi.org/10.1016/j.micres.2020.126504>
- Nipun, B., Dholaria, M., & Tailor, A. J. (2024). *Metagenomic Exploration Microbial Diversity of Industrial Effluent Contaminated Soils of Silvassa, India*. 9(6), 1232–1250.
- Nipun Mayuri Dholaria, B. (2024). Metagenomic Approaches for Studying Microbial Community of Soil: Review. *International Journal of Science and Research (IJSR)*, 13(2), 1048–1062. <https://doi.org/10.21275/SR24212114822>
- Nwachukwu, B. C., & Babalola, O. O. (2022). Metagenomics: A Tool for Exploring Key Microbiome With the Potentials for Improving Sustainable Agriculture. *Frontiers in Sustainable Food Systems*, 6. <https://doi.org/10.3389/fsufs.2022.886987>
- Pandit, P., Kumar, R., Kumar, D., Patel, Z., Pandya, L., & Joshi, C. (2020). *Microbial Genes Involved in the Aromatic Compound Degradation Deciphered Through Metagenomics Analysis of Industrial Wastewater*. <https://doi.org/10.21203/rs.3.rs-62394/v1>
- Patel, V., Sharma, A., Lal, R., Al-Dhabi, N. A., & Madamwar, D. (2016). Response and resilience of soil microbial communities inhabiting in edible oil stress/contamination from industrial estates. *BMC Microbiology*, 16(1). <https://doi.org/10.1186/s12866-016-0669-8>
- Pathak, A., Jaswal, R., Xu, X., White, J. R., Edwards, B., Hunt, J., Brooks, S., Rathore, R. S., Agarwal, M., & Chauhan, A. (2020). Characterization of Bacterial and Fungal Assemblages From Historically Contaminated Metalliferous Soils Using Metagenomics Coupled With Diffusion Chambers and Microbial Traps. *Frontiers in Microbiology*, 11. <https://doi.org/10.3389/fmicb.2020.01024>
- Pérez-De-Mora, A., Burgos, P., Madejón, E., Cabrera, F., Jaekel, P., & Schlöter, M. (2006). Microbial community structure and function in a soil contaminated by heavy metals: Effects of plant growth and different amendments. *Soil Biology and Biochemistry*, 38(2), 327–341. <https://doi.org/10.1016/j.soilbio.2005.05.010>
- Rahmeh, R., Akbar, A., Kumar, V., Al-Mansour, H., Kishk, M., Ahmed, N., Al-Shamali, M., Boota, A., Al-Ballam, Z., Shajan, A., & Al-Okla, N. (2021). Insights into Bacterial Community Involved in Bioremediation of Aged Oil-Contaminated Soil in Arid Environment. *Evolutionary Bioinformatics*, 17. <https://doi.org/10.1177/11769343211016887>
- Reddington, K., Eccles, D., O’Grady, J., Drown, D. M., Hansen, L. H., Nielsen, T. K., Ducluzeau, A. L., Leggett, R. M., Heavens, D., Peel, N., Snutch, T. P., Bayega, A., Oikonomopoulos, S., Ragoussis, I., Barry, T., Van Der Helm, E., Jolic, D., Richardson, H., Jansen, H., ... Brown, B. L. (2020). Metagenomic analysis of planktonic riverine microbial consortia using nanopore sequencing reveals insight into river microbe taxonomy and function. *GigaScience*, 9(6), 1–12. <https://doi.org/10.1093/gigascience/giaa053>
- Rekadwad, B. N., Shouche, Y. S., & Jangid, K. (2024). Community characterization of soil from the Central Deccan Plateau dry tropical deciduous forest (Ceddtruf) in central India using 16S rRNA gene amplicon sequencing. *Microbiology Resource Announcements*, 13(2). <https://doi.org/10.1128/mra.01134-23>
- Sandhu, M., & Paul, A. T. (2021). *Metagenomic Analysis for Taxonomic and Functional Potential of Polyaromatic Hydrocarbons (PAHs) and Polychlorinated Biphenyl (PCB) Degrading Bacterial Communities in Steel Industrial Soil*. 1–26. <https://doi.org/10.1371/journal.pone.0266808>
- Sandhu, M., Paul, A. T., & Jha, P. N. (2022). Metagenomic analysis for taxonomic and functional potential of Polyaromatic hydrocarbons (PAHs) and Polychlorinated biphenyl (PCB) degrading bacterial communities in steel industrial soil. *PLOS ONE*, 17(4), e0266808. <https://doi.org/10.1371/journal.pone.0266808>
- Singh, T., Pal, D. B., Bhatiya, A. K., Mishra, P. K.,



- Hashem, A., Alqarawi, A. A., AbdAllah, E. F., Gupta, V. K., & Srivastava, N. (2022). Integrated process approach for degradation of p-cresol pollutant under photocatalytic reactor using activated carbon/TiO₂ nanocomposite: application in wastewater treatment. *Environmental Science and Pollution Research*, 29(41), 61811–61820. <https://doi.org/10.1007/s11356-021-15454-5>
- Sousa, J., Silvério, S. C., Costa, A. M. A., & Rodrigues, L. R. (2022). Metagenomic Approaches as a Tool to Unravel Promising Biocatalysts from Natural Resources: Soil and Water. *Catalysts*, 12(4), 385. <https://doi.org/10.3390/catal12040385>
- Su, C., Xie, R., Liu, D., Liu, Y., & Liang, R. (2023). Ecological Responses of Soil Microbial Communities to Heavy Metal Stress in a Coal-Based Industrial Region in China. *Microorganisms*, 11(6), 1392. <https://doi.org/10.3390/microorganisms11061392>
- Su, H., Kuang, X., Ren, Y., & Luo, L. (2022). Biostimulants promote biodegradation of n-hexadecane by *Raoultella planticola*: Generation of lipopeptide biosurfactants. *Journal of Environmental Chemical Engineering*, 10(5). <https://doi.org/10.1016/j.jece.2022.108382>
- Sun, S., Wang, Y., Xu, C., Qiao, C., Chen, S., Zhao, C., Liu, Q., & Zhang, X. (2023). Reconstruction of microbiome and functionality accelerated crude oil biodegradation of 2,4-DCP-oil-contaminated soil systems using composite microbial agent B-Cl. *Journal of Hazardous Materials*, 447, 130808. <https://doi.org/10.1016/j.jhazmat.2023.130808>
- Thomas, J. C., Oladeinde, A., Kieran, T. J., Finger, J. W., Bayona-Vásquez, N. J., Cartee, J. C., Beasley, J. C., Seaman, J. C., McArthur, J. V., Rhodes, O. E., & Glenn, T. C. (2020). Co-occurrence of antibiotic, biocide, and heavy metal resistance genes in bacteria from metal and radionuclide contaminated soils at the Savannah River Site. *Microbial Biotechnology*, 13(4), 1179–1200. <https://doi.org/10.1111/1751-7915.13578>
- Vasanthrao, R., Nidhin, I. K., Taj, Z., & Chattopadhyay, I. (2025). Comprehensive whole metagenomics analysis uncovers microbial community and resistome variability across anthropogenically contaminated soils in urban and suburban areas of Tamil Nadu, India. *Frontiers in Microbiology*, 16(October), 1–29. <https://doi.org/10.3389/fmicb.2025.1649872>
- Vasudeva, G., Singh, H., Paliwal, S., & Pinnaka, A. K. (2022). Metagenomics: An Approach for Unraveling the Community Structure and Functional Potential of Activated Sludge of a Common Effluent Treatment Plant. *Frontiers in Microbiology*, 13. <https://doi.org/10.3389/fmicb.2022.933373>
- Wei, C., Sun, D., Yuan, W., Li, L., Dai, C., Chen, Z., Zeng, X., Wang, S., Tang, Y., Jiang, S., Wu, Z., Peng, X., Jiang, L., & Peng, S. (2021). *Metagenomics Revealing Molecular Profiling of Microbial Community Structure and Metabolic Capacity In Bamucuo, Tibet*. <https://doi.org/10.21203/rs.3.rs-505014/v1>
- Xu, R., Sun, X., Häggblom, M. M., Dong, Y., Zhang, M., Yang, Z., Xiao, E., Xiao, T., Gao, P., Li, B., & Sun, W. (2022). Metabolic potentials of members of the class Acidobacteriia in metal-contaminated soils revealed by metagenomic analysis. *Environmental Microbiology*, 24(2), 803–818. <https://doi.org/10.1111/1462-2920.15612>
- Yadav, M., Kumar, T., Kanakan, A., Maurya, R., Pandey, R., & Chauhan, N. S. (2022). Isolation and Characterization of Human Intestinal Bacteria *Cytobacillus oceanisediminis* NB2 for Probiotic Potential. *Frontiers in Microbiology*, 13. <https://doi.org/10.3389/fmicb.2022.932795>
- Zhang, X., Wu, K., Han, Z., Chen, Z., Liu, Z., Sun, Z., Shao, L., Zhao, Z., & Zhou, L. (2022). Microbial diversity and biogeochemical cycling potential in deep-sea sediments associated with seamount, trench, and cold seep ecosystems. *Frontiers in Microbiology*, 13. <https://doi.org/10.3389/fmicb.2022.1029564>