Metagenomic Analysis for Taxonomic and Functional Potential of Polyaromatic Hydrocarbons (PAHs) and Polychlorinated Biphenyl (PCB) Degrading Bacterial Communities in Steel Industrial Soil

Monika Sandhu
Birla Institute of Technology and Science Pilani: Birla Institute of Technology and Science
Atish T. Paul
Birla Institute of Technology and Science Pilani campus, Pilani
Prabhat Nath Jha (prabhatn.jha@gmail.com)
Birla Institute of Technology and Science

Research Article

Keywords: Metagenomic, Persistant organic pollutants, Biodegradation, Microbial community

Posted Date: June 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-564700/v1

Abstract

Iron and steel industries are the major contributors to persistent organic pollutants (POPs). The microbial community present at such sites has potential to remediate these contaminants. The present study highlights the metabolic potential of resident microbial community of PAHs and PCB contaminated soil nearby Bhilai steel plant, Chhattisgarh (India). The GC-MS/MS analysis of soil samples MGB-2 (sludge) and MGB-3 (dry soil) resulted in identification of different classes of POPs including PAHs {benzo[a]anthracene (nd; 17.69%), fluorene (15.89%, nd), pyrene (nd; 18.7%), benzo(b)fluoranthene (3.03%, nd), benzo(k)fluoranthene (11.29%; nd), perylene (5.23%; nd)} and PCBs (PCB-15, PCB-95, and PCB-136). Whole-genome metagenomic analysis by Oxford Nanopore GridION Technology revealed *Proteobacteria* (44.3%; 50.0%) to be the prominent phylum followed by *Actinobacteria* (22.1%; 19.5%) in MBG-2 and MBG-3, respectively. The sample MGB-3 was richer in terms of macronutrients (C, N, P) supporting high microbial diversity than MGB-2. Taxonomic *vis-à-vis* functional analysis identified *Burkholderia*, *Bradyrhizobium*, *Mycobacterium*, and *Rhodopseudomonas* as the keystone degrader of PAH and PCB. Overall, our results revealed the importance of metagenomic together with physicochemical analysis of contaminated site which improves the understanding of metabolic potential and adaptation of bacteria growing under stressful environment.

Introduction

Persistent organic pollutants (POPs) are anthropogenic chemicals that are enlisted in priority environmental pollutants due to their toxicity and persistence in the environment for a prolonged period (Jones and de Voogt 1998). PAHs/PCBs are strongly lipophilic and hence they easily enter the food chains. These characteristics are important since they are responsible for the detrimental effect on the environment and induce health threats to plants, animals, and humans (Engraff et al. 2011). Increased industrialization has led to the extensive production of such POPs which are also emitted during the production of steel (Zhang et al. 2015). The rise in these pollutant have led to adverse health and the environment effects, which inturn has resulted in extensive studies on the remediation of contaminated soil. Various physical and chemical technologies including chemical oxidation, electrokinetic remediation, solvent extraction, photocatalytic degradation, and thermal treatment are widely applied in remediation (Van Gerven et al. 2004). However, most of these treatment methods are unsustainable, disruptive and carry these PCBs/PAHs to the environment. Therefore, the utilization of existing contaminated soil-based bacterial communities can prove to be an alternative strategy for effective and viable degradation of POPs (Lu et al. 2019) as it has comparatively less technical hindrances than other remediation technologies.

A series of studies have been performed by the culture-dependent approach to isolate the most efficient biodegrader from such polluted sites (Issac et al. 2013). The contaminated soil environment consists of the genetic, species, and metabolic diversity of microbial biodegraders. Using a culture-dependent method only a minor fraction of POPs degrading bacteria can be obtained. Furthermore, it has been reported that enrichment of these cultures under lab conditions is less efficient in biodegradation than indigenous bacteria present in the contaminated soil (Isaac et al. 2015). To date, information related to taxonomic and functional interaction amongst the microbial communities that occur during the biodegradation process within the contaminated environment is skewed. The recent development of powerful culture-independent metagenomic approaches and the advancement of next-generation sequencing (NGS) technology provides a comprehensive insight into the total microbial community inhabiting contaminated sites and their metabolic capabilities. Several metagenomics studies conducted on PAH/PCB contaminated soil samples (Zhang et al. 2019) have highlighted microbial interaction playing a pivotal part in bioremediation of these POPs. Metagenomic sequencing technology such as Oxford Nanopore Technology (ONT) is very sensitive and can detect very low abundant microbial members present in the metagenome that are otherwise missed.

The present study aimed to investigate and provide an insight into the bacterial community in soils contaminated with such as PAH, PCB and to correlate their functional characteristics with respect to the biodegradation pathways. The soil samples were collected from nearby regions of Bhilai steel plant (one of Asia's biggest steel plants) in Chhattisgarh, India. The PCBs congeners and PAHs have been reported to be present in the wastes sites of industrialized area of this steel plant, India (Singh et al. 2015). The major pollution source of steel industries includes sinter, coke, and the blast furnace (Jiun-Horng et al. 2007).

The present study is the first comprehensive report on microbial community and metabolic potential with respect to biodegradation pathways of soil collected from polluted site nearby region of Bhilai steel plant.

Materials And Methods

Study site and sampling

The soil samples were collected from 2 different sites i.e sludge site (MGB-2) and dry soil waste site (MGB-3) from the polluted area near Bhilai steel plant, Chhattisgarh (21.1915° N, 81.4041° E), in India (Fig 1). Soil samples were randomly collected in sterile containers from a depth of about 0 to 10 cm of two sampling sites. It was then transported on ice pack and were stored at 4 °C in the lab for analyses. Physicochemical parameters such as pH, electrical conductivity, organic C, N, P, Mg, K, Na, Cl, Ca, S, Zn, Fe, Cu, and Mn of MGB-2 and MGB-3 were estimated using the standard protocol at the National Horticultural Research and Development Foundation, Nasik, India.

Extraction and determination of PAH and PCB in sediments

5 g of each collected sample (dry weight) was added into 50 ml Milli Q (MQ) water and was homogenized by vortexing for 15 min. After allowing it to stand for 30 min, 10 ml of acetone and hexane (1:1; *v/v*) were added to the falcon and was vortexed for 3 min. 2g NaCl was added and shaken vigorously for a few min (He et al. 2015). It was then centrifuged at 4000 x g for 5 min. The supernatant was subjected for further solid-phase extraction (SPE) of PCB and PAH using bond elute cartridge as per manufacturer's instruction (Agilent technologies, USA). The cartridge was conditioned with 2 ml acetonitrile for 2 min at 1000 x g followed by repeated conditioning with 2 ml MQ water. Extracted sample (2 ml) was loaded onto the cartridge and centrifuged for 5 min at 4000 x g. This step was repeated twice. Further, the sample elution was performed with methanol and hexane 1:1 (*v/v*) in 5 ml MQ by centrifuging for 2 min at 1000 x g. The

final elute was then collected through a Polytetrafluoroethylene (PTFE) filter (0.22 µ) in a separate vial and adjusted to 1 ml with nitrogen (Weiland-Brauer et al. 2017). GCMS-TQ8040 (Shimadzu, Japan) fitted with Scan/SIM was used for the qualitative analysis of PAHs and PCBs that are potentially present in the MGB-2 and MGB-3 samples. GC-MS/MS fitted with Flame ionization detector (FID) and an RTX-5 column (30 m × 0.32 mm × 0.25 µm) was used for analysis. GC conditions were set at 40°C with a 2 min hold and 10 °C/min increment to 80 °C, then 6 °C/min to 225 °C with 10 min hold. The presence of PCB was detected through SIM mode of GC-MS/MS.

Metagenome Sequencing and analysis

DNA extraction and processing for metagenome

Soil samples MGB-2 and MGB-3 were collected in triplicate and pooled together for each sample. DNA extraction was done using Powersoil® DNA Isolation Kit (Qiagen, USA) following the manufacturer's instructions. The metagenomic DNA sample was checked for integrity by agarose gel (1%) using a BioRad Gel documentation system and was quantified by Qubit 3.0 Fluorometer (Invitrogen, USA).

Preparation of library and whole metagenome sequencing

Metagenomic DNA from MGB-2 and MGB-3 were end-repaired using NEBnext ultra II kit (New England Biolabs, USA), cleaned up with 1x AmPure beads (Beckmann Coulter, USA). Native barcode ligation was performed with NEB blunt/TA ligase using NBD103 and cleaned with 1x AmPure beads. Qubit quantified barcode ligated DNA samples were pooled at an equimolar concentration to attain a 1 µg pooled sample. Adapter ligation (BAM), cleaning of library mix and elution of sequencing library were done as per the protocol by Gutierrez et al, 2018 and futher were used for whole-genome sequencing. The whole-genome library was prepared by using a Native Barcoding kit (EXP-NBD103). Barcode sequences are detailed in the supporting information (Table S1). The sequencing was performed using SpotON flow cell (R9.4) on MinKNOW 2.1 v18.05.5 with a 48 h sequencing protocol (Laver et al. 2015) on GridION X5 (Oxford Nanopore Technology (ONT, UK).

Data processing and analysis

The Nanopore raw reads (*'fast5'* format) were base-called (*'fastq5'* format) and demultiplexed using Albacore v2.3.1 and were uploaded to MG-RAST server (version 4.0.3) for taxonomic and functional analysis (Keegan et al. 2016). Functional annotation by SEED subsystems helps in predicting the abundance of genes assigned to metabolic pathways in soil. The sequenced reads were interpreted using multisource non-redundant ribosomal RNA database for taxonomic diversity and were determined using the contigLCA algorithm against the M5NR database for samples analyzed *via* whole-genome sequencing (WGS) (MG-RAST metagenome MGB-2 and MGB-3 identification numbers = mgm4822000.3, mgm4822001.3). The interpretation was based on E-value cut-off =1 x e^{-5} and sequence identity of 60% (Randle-Boggis et al. 2016; Brumfield et al. 2020).

Statistical analysis

Various alpha diversity indices were calculated to study species richness and evenness of the MGB-2 and MGB-3 using PAST4.03 software. Principal component analysis (PCA) plot was constructed using Bray-Curtis matrices with R studio v3.1.2. Comparison of samples MGB-2 and MGB-3 was done using Statistical Analysis of Metagenomic Profiles software (STAMP; Parks and Beiko, 2010) with a two-sided G-test (w/Yates'+ Fischer's). Comparative metagenome analysis was done mainly with RefSeq and SEED subsystem to obtain genus/functional abundance, respectively. Cytoscape software v3.7.1 was used to generate networking plots for study of the interaction of microbial communities of MGB-2 and MGB-3 involved in xenobiotic biodegradation pathways.

Results And Discussion

Physico-chemical analysis of the MGB-2 and MGB-3

Microbial community structure and function are determined by various environmental factors including nutritional status that in turn are governed by various physico-chemical parameters. Therefore, the physicochemical properties of the MGB-2 and MGB-3 were determined and are summarized in Table 1. The sample MGB-3 was richer in terms of macronutrients such as carbon (C), nitrogen (N), phosphorus (P) that greatly influence the composition of the microbial community (Luo et al. 2020). The organic carbon content, which represents the energy flow in the carbon cycle, was 0.85% (slightly high) in MGB-2 and 1.39% (high) in MGB-3 as compared to reference values. Our results indicated high carbon content in the given samples that could be because of aromatic organic hydrocarbons present in the contaminated soil. Industrial soil and effluent are considered to be a source of organic contaminants including POPs like PAHs (Cai et al. 2007). Because of the hydrophobic nature of these POPs, they tend to bind with the soil and hence add to the organic carbon content of the soil. The sample MGB-3 had very high nitrogen and phosphorus content (N, 734 kg ha^{-1}; P, 56.9 kg ha^{-1}) whereas MGB-2 had moderate nitrogen (430 kg ha^{-1}) and low phosphorus (17.66 kg ha^{-1}) content. MGB-3 exhibited high P and low C/P ratios indicating the possibility of higher microbial diversity as compared to the MGB-2. Several studies have confirmed that the soil with high microbial diversity has high P and low C/P ratios, while the environment with less P and high C/P ratios shows low microbial diversity (Delgado-Baquerizo et al. 2017). A higher level of these micronutrients in MGB-2 can be due to the high water content in the sample. Further, these results indicate that both MGB-2 and MGB-3 can support diverse microbial communities.

Overall, the level of micronutrients including K, Mg, Na, Mn was higher in MGB-2 than MGB-3. The level of K (1232 kg ha^{-1}) and Mg (768 kg ha^{-1}) was found to be higher than the reference value in MGB-2. The availability of inorganic nutrients serves for structural as well as catalytic functions. Therefore, it is supported with the fact that the total taxonomic and functional profile of the bacterial communities are predominantly driven by the availability of C and N and also by the presence of inorganic nutrients *i.e.*, Ca, K and Mg to some extent (Nicolitch et al. 2019). The result of the metal analysis indicated that the MGB-3 had comparatively higher metal content (Zn, Cu, and Fe) than the MGB-2 (Table 1). The result also indicated a high percent of Mn (44.12 mg kg^{-1})

in MGB-2 and Fe (14.98 mg kg^{-1}) in MGB-3, providing a suitable environment for microbes to undergo anaerobic biodegradation. The high metal content in both samples was due to the additives used in a steel factory. The presence of the metals in a soil sample can limit many microbial species whereas they can support the survival and growth of metal tolerant species. The presence of metals is in accordance with several reports that highlighted the presence of heavy metals along with organic contaminants such as PAH from iron and steel industrial soil sites (Bano et al. 2018). In addition, Mn(IV) and Fe(III) are known to act as terminal electron acceptors that efficiently remove aromatic compounds from the soil. Fe is the most widely found cofactor involved in deoxygenation reactions in biodegradation studies. It has been reported that Fe containing dioxygenases is incorporated into the active site either as iron center, Rieske [2Fe-2S] cluster or as heme prosthetic group during PAH and PCB biodegradation (Langenhoff et al. 1996).

Determination of PAH and PCB residues in contaminated soil samples

Cities with long industrial history contribute to the addition of organic pollutants in soil (Jiang et al. 2018; Ontiveros-Cuadras et al. 2019). Hence it was deemed fit to estimate the level of POPs mainly PAH and PCB in given soil samples by GC-MS/MS using Scan/SIM mode. GC-MS/MS triple quadrupole allows detection at very low (femtogram) limits in the matrix through the use of even greater selectivity with selected reaction monitoring (SIM) mode. Based on the data obtained from GC-MS/MS analysis, various PAHs along with PCB in industrial soil sample were identified. The structure and the relative percentage abundance of PAH identified in MGB-2 and MGB-3 are given in Fig 2. F (15.89%) and BkF (11.29%) were found to be dominant PAH species in MGB-2, while Pyr (18.7%) and BaA (17.69%) in MGB-3. High-molecular-weight (HMW; 4-6 rings) PAHs namely BcPhe (0.51%), BkF (11.29%), F (15.89%), IcdP (6.63%), Per (5.23%), and Tpl (4.70%), predominated in the MGB-2 sediments, which were not detected in MGB-3 (Fig 3). PAHs such as Indeno (1,2,3-cd) pyrene, benzo (k) fluoranthene, dibenzo (a,h) anthracene, chrysene, fluoranthene, acenaphthene, and fluorene have been reported to be in different operational units of steel industry (Khaparde et al. 2016).

In addition to PAHs, the MGB-2 and MGB-3 was also found to be contaminated with PCBs such as biphenyl, 1,1'-biphenyl, 4,4'-dichloro (PCB-15) in MGB-2 and biphenyl, 1,1'-biphenyl 2,2',3,',5,6 pentachloro (PCB-95) and 1,1'-biphenyl 2,2',3,3',6,6' hexachloro (PCB-136) in MGB-3 (Fig. S1-S2) (Aydin et al. 2014). It is well studied that the carcinogenic risk increases with the molecular weight or the aromatic ring of PAHs (Yang et al. 2002) and with the chlorine atoms in the case of PCBs. However, the data of source of PCBs has been scarce in industrial regions. The present results are in accordance with previous studies conducted by Patel et al. 2015 on Bhilai steel plant (Raipur, Chhattisgarh) soil, reporting the presence of PCBs ranging from of dichlorinated to hexachlorinated biphenyl in the sludge. From the data of organic contaminants, it appears that the high carbon content in MGB-2 and MGB-3 mentioned in the previous section could be co-related with the abundance of PAH and PCB in the soil sample.

Metagenomic analysis

Whole Genome Sequencing and assembly summary

The metagenomic approach provides a complete picture of biodegradation *vis-a-vis* microbes present within the environment and the functional genes involved in the bioremediation of contaminants. Whole-genome metagenomics studies are used not only for studying taxonomic diversity but also for elucidating metabolic pathways required for understanding pollutant degradation (Bouhajja et al. 2016). ONT has gained attention in recent times for community analysis due to its advantages over short-read sequencing technologies due to its ability of sequencing through repetitive regions within few minutes (Cummings et al. 2017). In the present study, the ONT platform for community analysis was used as it enables unbiased assembly of complete genome sequencing. Nanopore GridION X5 generated real-time, long-read, high-fidelity DNA sequence data. MG-RAST statistical analysis of dataset provided 275,844 sequences (totalling 539,360,072 bp; average length 1,955 bp) for MGB-2 and 193,221 (totalling 532,031,111 bp; average length 2,753 bp) for MGB-3. The downstream analyses of the total number of reads are detailed in Table S2. The datasets were used for various taxonomic, ecological, and functional analyses as described in previous section.

Analysis of sequence data for the extent of microbial diversity

The horizontal rarefaction curve indicated the significant depth of sampling, hence, represents sufficient sample coverage for diversity analysis (Fig S3). MGB2 and MGB-3 comprised 1839 and 1884 species respectively, indicating higher species richness in MGB-3 that is also evident from the data of Chao-α. The various diversity indices depicted in Table 2 suggest equivalent overall diversity which considers both species richness and abundance. MGB-3 was more diverse as compared to MGB-2 as calculated by Chao-α and Shannon's diversity (Table 2). Higher species richness in MGB-3 may be due rich C, N, and P content which are a major determinant of microbial growth (Cleveland and Liptzin, 2007). Beta diversity among MGB-2 and MGB-3 based on Bray-Curtis dissimilarity (Fig 4) depicted the overall distribution pattern of bacterial communities in MGB-2 and MGB-3 samples. Principal coordinates PC1 showed 98.7%, bacterial communities' ordinated closer to one another across axes are more similar in the two sites (MGB-2 and MGB-3) than 1.71% bacterial communities.

Taxonomic diversity and total microbial community

ONT and MG-RAST database search provided detailed microbial taxonomy of MGB-2 and MGB-3. Major differences were noticeable at the domain level, where MG-RAST derived reads accounted for bacteria (97.4%; 97.5%) followed by Eukaryota (1.4%; 1.5%), archaea (1.2%; 0.9%) and virus (0.02%; 0.04%) in MGB-2 and MGB-3, respectively. Comparison of the most abundant bacteria between samples, distribution of bacteria at phylum, order, class, family, and genus level as depicted in Krona plots representing the relative abundance of different bacterial genera identified in the two samples (Fig 5a-b). Among bacteria, phylum *Proteobacteria* (43.3%; 50.0%) was the most predominant followed by *Actinobacteria* (22.1%; 19.5%) and *Firmicutes* (6.3%; 5.4%) in both MGB-2 and MGB-3 samples. The predominance of *Proteobacteria* and *Actinobacteria* in the soil sample indicated good biodegradation capability, as similar observations were found in contaminated sites of PAH, PCB, and other related pollutants (Lu et al. 2019). Also, the abundance of *Proteobacteria* reveals that the microbial community plays a vital role in photosynthesis, nitrogen fixation, sulfur, and phosphorus metabolism. Both the sample included all six classes of bacteria namely *Alphaproteobacteria* (18.6%; 22.4%), *Betaproteobacteria* (8.8%; 10.7%), *Deltaproteobacteria* (8.2%; 7.1%), *Epsilonproteobacteria* (0.3%; 0.3%),

Gammaproteobacteria (7.2%; 8.9%) and *Zetaproteobacteria* (0.04%; 0.03%) with difference in their relative abundance as observed by Ma et al. 2015 as well. At class level, *Actinobacteria (*26.2%) was observed to be dominant in MGB-2 while *Alphaproteobacteria* (22.1%) in MGB-3. STAMP analysis with G-test (w/Yates'+ Fischer's) two-way comparison showed the degree of closeness and difference across MGB-2 and MGB-3 at class level. The analysis showed close similarity in both metagenome, however they differ in relative abundance. Among the classes, *Actinomycetales* (85.7%; 88.4%) were found to be distinctly observed in MGB-2 and MGB-3 followed by *Solirubrobacterales* (8.6%; 6.7%).

At the genus level, the composition of the bacterial community in the two samples was similar but they differed in terms of their relative abundance. *Streptomyces* (3.7%) and *Candidatus Solibacter* (2.7%) were the most predominant genera in MGB-2, while in MGB-3 it was *Mycobacterium* (3.1%) and *Streptomyces* (3.0%). RefSeq annotations at highest taxonomic classification i.e., at genus level possess significant abundance level with $p < 1\ e^{-5}$ for *Streptomyces, Candidatus Koribacter, Gemmata, Conexibacter*, *Anaeromyxobacter* and *Nocardioides* in MGB-2, while *Hypomicrobium*, *Pseudomonas*, *Mycobacterium*, *Roseomonas,* and *Methylobacterium* with $p < 1\ e^{-5}$ in MGB-3. The correlation coefficient (r^2) indicated a linear relationship between MGB-2 and MGB-3 as depicted on the scatter plot (Fig S4). The dominance of *Mycobacterium* in MGB-3 can be correlated with the presence of pyrene in MGB-3, as it have been reported to be isolated from pyrene contaminated soil (Cheung and Kinkle, 2001). *Mycobacterium* has the ability to degrade pyrene. The presence of the dominant genera in samples suggests the potential for degrading organic pollutants by the members of these genera. For instance, *Anaeromyxobacter* is known to be capable of anaerobic respiration and possesses genes involved in reductive dechlorination processes at contaminated sites (Stanford et al. 2002). Similarly, members of *Mycobacterium*, *Bradyrhizobium*, *Burkholderia* (Goris et al. 2004; Rodrigues et al. 2006), and *Pseudomonas* (Wittich et al. 1999; Hatamian- Zarmi et al. 2009) are well known aromatic pollutant degraders in contaminated soil (Seo et al. 2009).

Within the metagenome of MGB-2 and MGB-3, *Archaeal* communities showed dominance of *Euryarchaeota* (76.7%; 81.3%) at phylum level followed by *Crenarchaeota* (13.7%; 13.2%) and *Thaumarchaeota* (8.7%; 4.8%), *Korarchaeota* (0.9%; 0.7%). At genus level relative abundance for methane-producing *Methanosarcina* (10.5%; 12.1%) *Methanoculleus* (3.7%; 3.2%) ammonia-oxidizing *Nitrosopumilus* (5.9%; 3.3%), Sulfur reducing bacteria like *Thermococcus* (4.3%; 4.7%) unclassified (derived from *Euryarchaeota*) (4.3%; 4.2%), *Pyrococcus* (3.9%; 3.3%), *Sulfolobus* (3.6%; 3.3%) dominated in both MGB-2 and MGB-3. However, two-way statistical comparison at *Archaea* genus level revealed *Nitrosopumilus* ($p = 1.24\ e^{-4}$) and *Cenarchaeum* ($p = 5.18\ e^{-3}$) in MGB-2 while *Methanobrevibacterium* (p = 0.012) in MGB-3 to predominate indicating more methanogenesis occurring in MGB-3 (Samson et al. 2019). The contaminated soil was reported to possess many members of *Methanosarcina*, *Halobacterium, Euryarchaeota,* and *Crenarchaeota* of uncultured genera. The diversity of *Archaea* is found to be higher within hydrocarbon degrading communities in contaminated environment than the non-contaminated counterpart. (Zhang et al. 2012). The relative abundance of *Archeal* communities was more in MGB-2 that can be associated with the higher soil moisture and POPs content.

Eukaryota microbial communities showed a predominance of phylum *Ascomycota* (20.5%; 23.6%), *Streptophyta* (18.5%, 17.0%) and unclassified (derived from *Eukaryota*) (12.1%; 12.2%) in MGB-2 and MGB-3. Several studies conducted on contaminated soils have revealed the presence of *Ascomycota as the* most dominant eukaryotic phylum, as the indigenous ascomycete can transform or remove the pollutants (Aranda, 2016).

Functional diversity and metabolic potential of MGB-2 and MGB-3

The SEED subsystem analysis in MG-RAST assigned reads based on various functions that identified 60% of the total function constituting metabolisms of amino acid, carbohydrate, energy, lipid, cofactors, vitamins, biosynthesis of glycan, polyketides, terpenoids, xenobiotic biodegradation, and biosynthesis of secondary metabolites. The remaining 40% functions include cellular processes, organismal systems, genetic & environmental processing, and human diseases.

General metabolism of microbial community

Metagenomic analysis revealed the coordination amongst the bacteria and different functional genes detected in metabolic pathways. Further, it is known that under stress conditions requires extra energy metabolism. Annotated genes involved in methane, C, N, and S metabolism pathways were identified. Predicted genes from the MGB-2 and MGB-3 metagenome were used as input in the KEGG mapper for functional annotation. Table 3a depicts a list of genes responsible for methane metabolism that are mainly attributed to methanotrophs and methanogens. It indicated the presence of hydrogenotrophic and methylotrophic methanogenesis pathways in both metagenomes. The presence of members belonging to *Nocardioides* (11%) and *Methylobacterium* (12%) were the major potential contributors of methanogenesis in MGB-2 and MGB-3, respectively. At the same time, the predominance of archaeal members belonging to *Methanosarcina*, *Methanoculleus,* and *Methanothermobacter* predominant in MGB-3 indicated more methane metabolism as compared to MGB-2.

Metagenome sequences analysed for nitrogen metabolism resulted in the identification of genes annotated for nitrate reductase (*napA, napB*), nitrite reductase (*nirB, nirD, nirK, nirS*), and nitric oxide reductase (*norB, norC, nosD*, *norZ)* in both the metagenome. Based on the number of reads, the dissimilatory N metabolism was noted to be more dominant than the assimilatory in both metagenome, which indicated cellular metabolism is more active than the cell synthesis process in the contaminated soil. The genome of the microbial communities also showed potential for the biological nitrogen fixation process which was evident from the presence of genes encoding the enzyme (*nifB, nifE nifK*, *nifN*) responsible for nitrogen fixation to ammonia. Bacteria playing a dominant role in N metabolism at genus level in this study are similar to that of industrial polluted sites reported by Devpura et al. 2017. The absence of nitrogen-fixing reads can also be related to the high nitrogen (biologically available) and presence of aerobic conditions in MGB-3. The relative abundance of N metabolism reads accounted for *Rhodopseudomonas* (11%) in MGB-2, whereas it was for *Mycobacterium* (14%) reads in MGB-3 at the genus level. These genes related to nitrogen fixation were assigned to nitrogen-fixing bacteria like *Bradyrhizobium*, *Frankia*, *Mesorhizobium*, *Sinorhizobium,* and *Rhizobium* in MGB-2 only, because of high water content. Also, Nitrite-oxidizing bacteria (NOB) and Ammonia oxidizing (AOB) that are responsible for conversion of NO_3 to ammonium namely *Nitrosomonas*, *Nitrospira,* and *Nitrobacter* were identified in MGB-2. However, no nitrogen fixation reads were identified in MGB-3. Genes encoding sulfite reductases belonging to the oxidoreductase family were found in both the metagenome. Assimilatory sulfate reduction involves oxidation of

sulfate to hydrogen sulfide through series of enzymes identified and are provided in Table 3b. *Anaeromyxobacter*, *Haliangium*, *Methylobacterium* were dominant in MGB-2, while *Mycobacterium*, *Conexibacter*, *Bordetella,* and *Pseudomonas* in MGB-3. Comparative analysis of both the metagenome based on their metabolism at SEED level 1 revealed significant dominance of methane metabolism ($p = 5.61\ e^{-4}$), citrate cycle ($p = 1.62\ e^{-3}$), carbon fixation ($p = 2.61\ e^{-3}$), pyruvate metabolism ($5.80\ e^{-3}$), aminobenzoate degradation (p = 0.031), selenocompound metabolism (p = 0.038) and benzoate degradation (p = 0.045) in both samples. However, lipoic acid metabolism ($p = 7.58\ e^{-4}$), drug metabolism ($p = 7.81\ e^{-3}$), folate biosynthesis (p = 0.17), fatty acid metabolism (p = 0.042) and lysine degradation (p = 0.042) were found to be comparatively high in abundance in MGB-3 (Fig 6a). It also revealed significant dominance of *Intrasporangum (*<$1\ e^{-5}$), *Rubrobacter* (<$1\ e^{-5}$), *Candidatus Koribacter* (<$1\ e^{-5}$), *Norcardiodes* ($1.77\ e^{-4}$), *Anaeromyxobacter* ($4.71\ e^{-4}$) in MGB-2 while *Mycobacterium* (<$1\ e^{-5}$), *Hypom*icrobium (<$1\ e^{-5}$), *Mesorhizobium* (<$1\ e^{-5}$), *Rhodobacter* (<$1\ e^{-5}$), *Methylobacterium* (<$1\ e^{-5}$), *Pseudomonas* ($1.65\ e^{-4}$) and *Burkholderia* ($3.33\ e^{-3}$) in MGB-3 (Fig 6b). The results indicated the active metabolism is represented by bacteria belonging to phylum *Actinobacteria* in MGB-2 while *Proteobacteria* in MGB-3.

The comparative analysis of gene sequences revealed abundance of P transporter in MGB-2 that can be correlated with the low P content than MGB-3. Various genes encoding proteins involved in phosphate-recycling mechanisms, such as *phnA, phnE, phnW* and *phnX*, (phosphonate transporters), *pstA, pstB, pstC*, and *pstS* (high-affinity phosphate transporters), and *phoR, phoA, phoP* and *phoD* (two-component systems) were detected across both the metagenome (Fig S5a). The availability number and abundance of reads for phosphate mechanism in low phosphate-phosphorous environment indicated these mechanism that helps them to cope up within such environment. Moreover, bacteria can survive due to the presence of resistant genes toward toxic metals such as copper, lead and nickel as a part of their defense mechanism, which is recruited further for cleaning the contaminated environments. The bioavailability of metals depends upon their interaction with the organic compounds present in the soil. The high abundance of Mn content in MGB-2 can be supported with the significantly high reads for Mn transporter in MGB-2 than MGB-3.

Bacteria survive by expressing several metal-resistant Cu(I)-responsive transcriptional regulator proteins. Genes corresponding to the above-mentioned functions were identified in the metagenome that leads to the accumulation of Cu(I) in an anoxic environment (Giachino A and Waldron KJ, 2020). Mn and Co efflux protein CorC were identified in both metagenomes. Cu homeostasis and tolerance genes namely *CutE* and *CutC* were found to be in abundance for *Bacillu*s, *Candidatus*, *Kineococcus*, *Mesorhizobium, Mycobacterium*, *Parvibaculum*, *Salinispora*, *Streptosporangium* in MGB-2, and *Nocardioides*, *Thiobacillus*, *Idiomarina*, *Rhodopseudomonas*, *Aromatoleum,* and *Bacillus* in MGB-3. The distinct relative abundance of iron acquisition and metabolism reads was observed for *Pseudomonas* (22%; 26%) in MGB-2 and MGB-3, respectively. Further investigation resulted in the identification of siderophore-producing bacteria *Pseudomonas* (43%; 51%) and *Streptomyces* (13%; 5%) in both samples. The comparative analysis reflected the higher dominance of iron metabolizing bacteria in MGB-3 than MGB-2 that is well matched to the data of soil analysis in terms of higher iron content in MGB-3 than MGB-2. The primary mechanism of Zn uptake involves ABC transporters such as *znuA, znuC,* and *znuR*. These Zn transporters present in both metagenome aids in transporting Zn across the membrane. Secondary mechanism of Zn uptake by cation transporters Zn tolerance reads at genus level was most dominant for *Anaeromyxobacter* (40%; 41%) in MGB-2 and MGB-3. Comparison between MGB-2 and MGB-3 of functional gene annotation using SEED subsystem for membrane transport revealed a significant level ($p > 0.05$) of abundance for Na^+/H^+ antiporters and Mn transporter MntH, Mn ABC transporter SitD, TadA, Zn ABC transporter ZnuA (Fig S5b-1d). MGB-3 showed higher Na^+/H^+ transporters indicating the exchange of the ions across the membrane to maintain homeostasis.

Abundance of Xenobiotics degradation and metabolism gene related to biodegradation

Overall, 708 (MGB-2) and 760 (MGB-3) annotated genes with varying levels of abundance (Fig S6) in two samples were identified corresponding to 17 pathways linked with xenobiotic biodegradation and metabolism. It was found that the annotated pathways for xenobiotic degradation and metabolism accounted for 2% of all 60 % metabolic functions (Fig 7a). The presence of these pathways is of great relevance as they are POPs enlisted by the U.S. Environmental Protection Agency. The presence of reads for biphenyl degradation, dioxin degradation, PAH pathways can be further correlated with the presence of PCB and PAH as detected in the MGB-2 and MGB-3 samples (Fig 7b). Further, chlorocyclohexane and chlorobenzene [PATH: Ko00361], and benzoate [PATH: Ko00362] degrading pathway were most prominent in MGB-2 (33.76%; 37.39%) and MGB-3 (39.55%; 33.96%). In addition to these degradative pathways, dioxin [PATH: Ko00621] and PAH [PATH: Ko00624] degradation pathways were also observed in MGB-2 (3.27%; 2.54%) and MGB-3 (4.29%; 2.80%) respectively. It is reported that chlorobenzene, PAH, PCB and benzoate are prominent contaminants of dye and steel industries.

The annotated enzymes hydroxyquinol 1,2-dioxygenase (*chqB*), pentachlorophenol 4-monooxygenase (*pcpB*), chorismate mutase/prephenate dehydratase (*pheA*), carboxymethylenebutenolidase (*clcD*), **haloacid** dehalogenase (*hadL*), benzene 1,2-dioxygenase/ toluene 1,2-dioxygenase (*bedC1/todC1*), biphenyl-2,3-diol 1,2-dioxygenase (*bphC*), catechol dioxygenase (*catA*), and muconate cycloisomerase (*catB*) were contributed mainly by members of genera in two communities suggesting coordinated degradation of chlorocyclohexane and chlorobenzene compounds depicted in Table 4a. Studies on chlorobenzene biodegradation highlighted that these compounds metabolized *via* the *ortho*-cleavage pathway (Mars et al., 1999) as well as by the *meta*-cleavage pathway (Mars et al., 1997).

Biphenyl degradation with annotated reads for *bphA, bphC, bphD, bphE,* and *bphF* were identified in both the metagenome. The abundance and diversity of the biphenyl/PCB enzyme and the corresponding bacterial genera in both the metagenome are listed in Table 4b. The annotation of *bph*A in both the metagenome MGB-2 and MGB-3 have indicated biphenyl degradation assigned to genus *Mycobacterium* in both the metagenome. In addition to *Mycobacterium*, it was also assigned to *Polaromonas* in MGB-3. Biphenyl 2, 3-diol, 1, 2-dioxygenase (*bph*C) is another important enzyme that was noted in both the metagenome. It has been assigned to *Mycobacterium* in MGB-2 while genus *Alkalilimnicola*, *Mycobacterium,* and *Polynucleobacter* in MGB-3. Aerobic and anaerobic degradation of PCB by bacteria have been studied in detail (Komancova et al. 2003; Jorge et al. 2006).

PAH degradation pathway was identified in MGB-2 and MGB-3 with annotated genes encoding enzymes required for degradation being listed in Table 4c. The presence of genes encoding hydroxychromene-2-carboxylate isomerase (*nahD*), naphthalene dioxygenase (*nahAc*), salicylate hydroxylase (*nahG*), and other

genes including *nidA, nidB, nidD, phdF, phdG, phdI,* and *phdJ* indicated complete pathway for degradation of PAH. The presence of reads for *nidA* gene was significantly correlated with the degradation of pyrene (Lu et al. 2019) that is in accordance with the presence of pyrene in metagenome MGB-3 detected by GC-MS/MS. It is reported that *nid*A gene is responsible for the synthesis of the large subunit of PAH dioxygenase involved in the degradation of PAHs such as phenanthrene, pyrene, benzo[a]pyrene, etc. (Pagnout et al. 2007). These enzymes are contributed by members of genera *Pseudoalteromonas, Aromatoleum*, *Dechloromonas*, *Agrobacterium*, *Mesorhizobium* in MGB-2, while *Mycobacterium*, *Parvibaculum*, *Ruegeria*, *Burkholderia*, *Aromatoleum*, *Bradyrhizobium* in MGB-3 communities suggesting synergistic degradation of these PAH.

Annotated gene/enzyme and assigned bacterial genera involved in benzoate degradation pathway identified in MGB-2 and MGB-3 are listed in table 4c. The presence of benzoate ligase in MGB-2 indicated anaerobic degradation of benzoate. Benzoate is the most common intermediate in the metabolism of aromatic compounds and in the anaerobic condition, it is converted into benzoyl-CoA by benzoate ligase. The presence of genes encoding anaerobic degradation *via* benzoyl coA (*badA, badD, badE, badF, badH, badI, hbaA, hbaB, hcrC*) and aerobic *via* protocatechuate (*lig*AB, *ligC, ligI, ligJ, pobA, pcaB, pcaC, pcaDF, pcaGH* and *pcaIJ*) catechol (*catE/ dmpB*) indicates complete pathway for degradation of benzoate. Protocatechuate can be mineralized *via* ortho- cleavage by protocatechuate 3,4-dioxygenase (PcaGH) and meta cleavage by protocatechuate 4,5-dioxygenase (LigAB). Anaerobic degradation *via* benzoyl-CoA has also been documented in a variety of facultative anaerobes, the denitrifying *Thauera* (Heider et al. 1998), *Aromatoleum*, *Azoarcus* (Valderrama et al. 2012), *Magnetospirillum* strains (Meyer-Cifuentes et al. 2017), and the photoheterotroph *Rhodopseudomonas* (Egland et al. 1995).These enzymes are contributed possibly by members of genera *Aromatoleum*, *Arthrobacter*, *Burkholderia*, *Bradyrhizobium*, *Cupriavidus*, *Magnetospirullum*, *Methylobacterium*, *Rhodomicrobium*, *Rhodopseudomonas,* and *Polaromonas* in both the metagenome communities suggesting synergistic degradation of benzoate. Comparison between MGB-2 and MGB-3 of functional gene annotation using SEED subsystem by STAMP for peripheral degradation revealed the abundance for, muconate cycloisomerase ring hydroxylating dioxygenase, naphthalene dioxygenase, acetaldehyde dehydrogenase, 2-hydroxy-6-oxo-6-phenyl hexa-2,4-dienoate hydrolase, and benzoate transport protein in MGB-2 while biphenyl 2,3 diol dioxygenase, benzaldehyde dehydrogenase, *nap/bph* dioxygenase, ortho-halobenzoate 1,2- dioxygenase, phenol dioxygenase, 1,2- dihydroxycyclohexa-3,5-diene-1-carboxylate dehydrogenate in MGB-3 (Fig.S5). Similarly, two-way comparison between MGB-2 and MGB-3 of functional gene annotation using SEED subsystem for aromatic degradation revealed significant level of abundance for benzoate ligase (p = 2.63 e^{-3}) and 4-hydroxyphenylacetate 3-monooxygenase (p = 8.93 e^{-3}) in MGB-2. The presence of 2-hydroxycyclohexanecarboxyl dehydrogenase with p = 2.39 e^{-3}, indicates anaerobic degradation of benzoate in MGB-3 as well. However, aerobic degradation of biphenyl was also found in MGB-3 as indicated by the abundance of biphenyl- 2,3-diol 1,2-dioxygenase with p = 0.035 in MGB-3 (Fig. S7-S8). Further, comparison between MGB-2 and MGB-3 at functional gene annotation using RefSeq for aromatic degradation revealed significant level of abundance of *Acidothermus* (p=2.68 e^{-3}), *Nakamurella* (p=6.39 e^{-3}) in MGB-2 while *Mycobacterium* (p = 3.30 e^{-3}), *Leptothrix* (p=7.89 e^{-3}) and *Polynuclearbacter* (p=0.034) in MGB-3 (Fig. S9).

Our findings suggest that biphenyl (PCB) and pyrene/ phenanthrene (PAH) biodegradation pathways could be linked together *via* common intermediate protocatechuate pathway and undergoes complete degradation through the common protocatechuate branch of the β-ketoadipate pathway. Therefore, reconstruction of complete biphenyl/PCB and PAH degradative pathways based on the annotated genes was done and shown in Fig 8. Cytoscape-based networking revealed microbial interaction in the xenobiotic biodegradation pathway. The key biodegraders in xenobiotic biodegradation pathways (aromatic halogenated, dioxin, PAH, PCB/biphenyl, catechol, protocatechuate, benzoyl-CoA) were found to be *Arthrobacter chlorophenolicus*, *Meiothermus ruber*, *Cupriavidus metallidurans*, *Burkholderia xenovorans*, *Rubrobacter xylanophilus*, *Bradyrhizobium* sp. BTAi1, *Bradyrhizobium japonicum*, *Sphingobium japonicum*, *Pseudomonas aeruginosa*, *Polaromonas* sp. JS666 and *Rhizobium leguminosarum* in MGB-2 (Fig 9). In MGB-3 it was *Bradyrhizobium japonicum*, *Rhizobium leguminosarum*, *Mycobacterium* sp. KMS, *Polaromonas* sp. JS666, *Rhodopseudomonas palustris*, *Xanthobacter autotrophicus*, *Mycobacterium smegmatis*, *Burkholderia cenocepacia,* and *Burkholderia xenovorans* (Fig 10). The greater abundance of these genes and genera in sample MGB-2 compared to sample MGB-3 suggests that there higher degrading capacity in sample MGB-2. However, the key biodegrader were found to be *Bradyrhizobium, Burkholderia, Mycobacterium* and *Rhodopseudomonas* in both the metagenome.

Conclusions

The present metagenomic study highlighted microbial function annotation, extensive degradation capabilities in terms of xenobiotic degradation pathways and correlated with the presence of PAH and PCB in the contaminated soil of steel plant. In addition, physico-chemical profiling of the soil samples provided valuable information regarding the presence of organic (C/N/P), inorganic nutrient (Ca, K, Mg, Na, Mn) and metal (Fe, Mn, Cu, Zn) present, that can be important parameter for designing biodegradation strategies. Higher proportions of *Proteobacteria* and *Actinobacteria* indicated the two sample possess good biodegradation potential. Moreover, the coordination among different biodegraders along with presence of functional genes involved in biodegradation pathways and energy metabolism has provided in-depth understanding regarding their survival under stress conditions of persistent organic pollutants. Moreover, these potential biodegrader such as *Bradyrhizobium, Burkholderia, MycobacteriumRhodopseudomonas* and *Pseudomonas* identified in the present study can be further selected and further could be exploited for enhancing bioremediation. Therefore, it can be concluded that investigating microbial community and exploring their potential for biodegradation is a critical factor in maximizing the efficacy of the bioremediation process.

Declarations

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files

Acknowledgements

Our special thanks to Central instrumentation Facility, Birla Institute of Technology and Sciences (BITS), Pilani campus for providing GC-MS/MS facility. MS acknowledges financial assistance provided by DST, India.

Funding

Authors are thankful to Department of Science and Technology, New Delhi, India for financial assistance under Women Scientist Scheme B [**SR/WOS-B/570/2016**] to MS.

Author information

Affiliations

Department of Biological Sciences, Birla Institute of Technology and Science Pilani, Pilani campus, Pilani-333031, Rajasthan, INDIA

Monika Sandhu & Prabhat N. Jha

Department of Pharmacy, Birla Institute of Technology and Science Pilani, Pilani campus, Pilani-333031, Rajasthan, INDIA

Atish T. Paul

Contributions

M.S. and P.N.J. designed the project. A.T.P. designed GC-MS/MS based experiment and analyzed respective data. M.S. conducted all experiments. M.S. and P.N.J. analyzed and interpreted the results, and wrote the manuscript. All of the authors contributed to finalizing the manuscript.

Correspondencing Author

Correspondence to Prabhat N. Jha.

Ethics declarations

Competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

References

Aranda E (2016) Promising approaches towards biotransformation of polycyclic aromatic hydrocarbons with *Ascomycota* fungi. Curr Opin Biotechnol 38: 1-8

Aydin YM, Kara M, Dumanoglu Y, Odabasi M, Elbir T (2014) Source apportionment of Polycyclic Aromatic Hydrocarbons (PAHs) and Polychlorinated Biphenyls (PCBs) in ambient air of an industrial region in Turkey. J Atmos Env 97:271-285

Bano S, Pervez S, Chow JC, Matawle JL, Watson JG, Sahu RK, Srivastava A, Tiwari S, Pervez YF, Deb MK (2018) Coarse particle (PM10–2.5) source profiles for emissions from domestic cooking and industrial process in Central India. J Sci Tot Env 627:1137–1145

Bouhajja E, Agathos SN, George IF (2016) Metagenomics: Probing pollutant fate in natural and engineered ecosystems. Biotechnol Adv 34(8):1413-1426. https://doi.org/ 10.1016/j.biotechadv.2016.10.006.

Brumfield KD, Huq A, Colwell RR, Olds JL, Leddy MB (2020) Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. PLoS ONE 15 (2):e0228899. https://doi.org/ 10.1371/journal.pone.0228899.

Cai Y, Mo H, Wu T, Zeng Y, Katsoyiannis A, Ferard F (2007) Bioremediation of polycyclic aromatic hydrocarbons (PAHs)-contaminated sewage sludge by different composting processes. J Hazard Mater 142 (1-2):535-542. https://doi.org /10.1016/j.jhazmat.2006.08.062.

Cheung PY, Kinkle BK (2001) *Mycobacterium* diversity and pyrene mineralization in petroleum-contaminated soils. Appl Environ Microbiol 67(5): 2222–2229. https://doi.org/ 10.1128/AEM.67.5.2222-2229.2001.

Cleveland CC, Liptzin D (2007) C:N:P stoichiometry in soil: is there a “Redfield ratio” for the microbial biomass? Biogeochemistry. 85:235-252. https://doi.org/10.1007/s10533-007-9132-0.

Cummings PJ, Olszewicz J, Obom KM (2017) Nanopore DNA Sequencing for Metagenomic Soil Analysis. J Vis Exp 130:55979. doi: 10.3791/55979.

Delgado-Baquerizo M, Eldridge DJ, Ochoa V, Gozalo B, Singh BK, Maestre FT (2017) Soil microbial communities drive the resistance of ecosystem multi functionality to global change in drylands across the globe. Ecol Lett 20:1295–1305. https://doi.org/ 10.1111/ele.12826.

Devpura N, Jain K, Patel A, Joshi CG, Madamwar D (2017) Metabolic potential and taxonomic assessment of bacterial community of an environment to chronic industrial discharge. Int Biodeterior Biodegradation 123:216-227. https://doi.org/10.1016/j.ibiod.2017.06.011.

Egland PG, Gibson J, Harwood CS (1995) Benzoate-coenzyme A ligase, encoded by *bad*A, is one of three ligases able to catalyze benzoyl-coenzyme A formation during anaerobic growth of *Rhodopseudomonas palustris* on benzoate. J Bacteriol 177 (22):6545-6551 https://doi.org/ 10.1128/jb.177.22.6545-6551.

Engraff M, Solere C, Smith KEC, Mayer P, Dahllof I (2011) Aquatic toxicity of PAHs and PAH mixtures at saturation to benthic amphi pods: linking toxic effects to chemical activity. Aquat Toxicol 102:142–149

Galazka A, Grzadziel J, Gałazka R, Ukalska-Jaruga A, Strzelecka J, Smreczak B (2018) Genetic and Functional Diversity of Bacterial Microbiome in Soils with Long Term Impacts of Petroleum Hydrocarbons. Front Microbiol 9:1923 https://doi.org/10.3389/fmicb.2018.01923.

Giachino A, Waldron KJ (2020) Copper tolerance in bacteria requires the activation of multiple accessory pathways. Mol Microbiol 114 (3):377-390. https://doi.org/10.1111/mmi.14522.

Godheja J, Shekhar SK, Satyanarayan GNV, Singh SP, Modi DR (2017) Antibiotic and heavy metal tolerance of some indigenous bacteria isolated from petroleum contaminated soil sediments with a study of their aromatic hydrocarbon degradation potential. Int J Curr Microbiol App Sci 6(3):194-211

Goris J, De Vos, P, Caballero-Mellado J, Park J, Falsen E, JF, Tiedje JM, Vandamme P (2004) Classification of the biphenyl- and polychlorinated biphenyl-degrading strain LB400^{T} and relatives as *Burkholderia xenovorans* sp. nov. Int J Syst Evol Microbiol 54 (5): 21

Grostern A, Sales CM, Zhuang WQ, Erbilgin O, Alvarez-Cohen L (2012) Glyoxylate metabolism is a key feature of the metabolic degradation of 1,4-dioxane by *Pseudonocardia dioxanivorans* strain CB1190. Appl Environ Microbiol 78 (9): 3298–3308. https://doi.org/ 10.1128/AEM.00067-12.

Guarino C, Zuzolo D, Marziano M, Conte B, Baiamonte G, Morra L, Benotti D, Gresia D, Stacul E, Cicchella D, Sciarrillo R (2019) Investigation and Assessment for an effective approach to the reclamation of Polycyclic Aromatic Hydrocarbon (PAHs) contaminated site: SIN Bagnoli, Italy. Scientific Reports 9:11522. https:doi.org/ 10.1038/s41598-019-48005-7.

Harwood, CS, Parales RE (1996) The beta-ketoadipate pathway and the biology of self-identity. Annu Rev Microbiol 50:553–590. https://doi.org/ 10.1146/annurev.micro.50.1.553.

Hatamian-Zarmi A, Shojaosadati SA, Vasheghani-Farahani E, Hosseinkhani S, Emamzadeh A (2009) Extensive biodegradation of highly chlorinated biphenyls and Aroclor 1242 by *Pseudomonas aeruginosa* TMU56 isolated from contaminated soils. Int Biodeter Biodegr 63:788–794

He Q, Sanford RA (2003) Characterization of Fe (III) reduction by chlororespiring Anaeromyxobacter dehalogens.. Appl Environ Microbiol 69:2712-2718. https://doi.org/10.1128/AEM.69.5.2712-2718.2003.

He Z, Wang L, Peng Y, Luo M, Wang W, Liu X (2015) Determination of selected polychlorinated biphenyls in soil and earthworm (*Eisenia fetida*) using a QuEChERS-based method and gas chromatography with tandem MS. J Sep Sci 38 (21):3766-3773. https://doi.org/ 10.1002/jssc.201500624.

Heider J, Boll M, Breese K, et al (1998) Differential induction of enzymes involved in anaerobic metabolism of aromatic compounds in the denitrifying bacterium *Thauera aromatica*. Arch Microbiol 170:120–131. https://doi.org/10.1007/s002030050623.

Isaac P, Sanchez LA, Bourguignon N, Cabral ME, Ferrero MA (2013) Indigenous PAH-degrading bacteria from oil-polluted sediments in Caleta Cordova, Patagonia Argentina. Int Biodeterior Biodegradation 82:207-214. https://doi.org/10.1016/j.ibiod.2013.03.009.

Isaac P, Martinez FL, Bourguignon N, Sanchez LA, Ferrero MA (2015) Improved PAHs removal performance by a defined bacterial consortium of indigenous *Pseudomonas* and *Actinobacteria* from Patagonia, Argentina. Int Biodeterior Biodegrad 101:23-31. https://doi.org/ 10.1016/j.ibiod.2015.03.014.

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. Environ Sci Technol 52 (15):8558-8567. https://doi.org/10.1021/acs.est.7b06647.

Jiun-Hornga T, Kuo-Hsiungb L, Chih-Yub C, Jian-Yuanb D, Ching-Guanb C, Hung-Lungc C (2007) Chemical constituents in particulate emissions from an integrated iron and steel facility. J Haz Mat 147:111–119. http://doi.org/10.1016/j.jhazmat.2006.12.054.

Jones KC, de P Voogt P (1998) Persistent organic pollutants (POPs): state of the science. Environ Pollut 100:209-221

Jorge LM, Rodrigues CAK, Michael RA, John FQ, Olga VM, Tamara VT, James MT (2006) Degradation of Aroclor 1242 Dechlorination Products in Sediments by *Burkholderia xenovorans* LB400 (*ohb*) and *Rhodococcus* sp. Strain RHA1(*fcb*). Appl Environ Microbiol 72 (4):2476-2482. https://doi.org/10.1128/AEM.72.4.2476-2482.2006.

Keegan KP, Glass EM, Meyer F (2016) MG-RAST, a metagenomics service for analysis of microbial community structure and function. Methods Mol Biol 207–233. https://doi.org/ 10.1007/978-1-4939-3369-3_13.

Khaparde VV, Bhanarkar AD, Majumdar D, Rao CVC (2016) Characterization of polycyclic aromatic hydrocarbons in fugitive PM10 emissions from an integrated iron and steel plant. J Sci Tot Env 562:155–163. https://doi.org/ 10.1016/j.scitotenv.2016.03.153.

Komancova M, Jurcova I, Kochankov L, Burkhard J (2003) Metabolic pathways of polychlorinated biphenyls degradation by *Pseudomonas* sp. 2. Chemosphere 50 (4): 537-543. https://doi.org/ 10.1016/S0045-6535(02)00374-0.

Langenhoff AAM, Zehnder AJB, Schraa G (1996) Behaviour of toluene, benzene and naphthalene under anaerobic conditions in sediment columns. Biodegradation 7:267–274. https://doi.org/ 10.1007/BF00058186.

Laver T, Harrison J, O'Neill PA, Moore K, Farbos A, Paszkiewicz K, Studholme DJ (2015) Assessing the performance of the Oxford Nanopore Technologies MinION. Biomol Detect Quantif 3:1–8. https://doi.org/ 10.1016/J.BDQ.2015.02.001.

Lovley DR, Ueki T, Zhang T, et al (2011) Geobacter: the microbe electric's physiology, ecology, and practical applications. Advances Microbial Physiology 59:1-100. https://doi.org/10.1016/b978-0-12-387661-4.00004-5.

Lu C, Hong Y, Liu J, Gao Y, Ma Z, Yang B, Ling W, Waigi MG (2019) A PAH-degrading bacterial community enriched with contaminated agricultural soil and its utility for microbial bioremediation. J Env Pol 251:773-782. https://doi.org/10.1016/j.envpol.2019.05.044.

Ma Q, QuY, Shen W, et al (2015) Bacterial community compositions of coking wastewater treatment plants in steel industry revealed by Illumina high-throughput sequencing. J Bior Tech 179:436–443. https://doi.org/ 10.1016/j.biortech.2014.12.041.

Mars AE, Kasberg T, Kaschabek SR, van Agteren, MH, Janssen DB, Reineke W (1997) Microbial degradation of chloroaromatics: use of the meta-cleavage pathway for mineralization of chlorobenzene. J Bact 179:4530-4537

Mars AE, Kingma J, Kaschabek SR, Reineke W, Janssen DB (1999) Conversion of 3-chlorocatechol by various catechol 2,3-dioxygenases and sequence analysis of the chlorocatechol dioxygenase region of *Pseudomonas putida* GJ31. J Bact 181:1309-1319

McLeod MP, Warren RL, et al (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA 103:15582–15587

Meyer-Cifuentes I, Martinez-Lavanchy PM, Marin-Cevada V, Bohnke S, Harms H, Muller, JA, Heipieper, HJ (2017) Isolation and characterization of *Magnetospirillum* sp. strain 15-1 as a representative anaerobic toluene-degrader from a constructed wetland model. PLoS One. 12(4). https://doi.org/10.1371/journal.pone.0174750.

Nicolitch O, Feucherolles M, Churin JL, et al (2019) A microcosm approach highlights the response of soil mineral weathering bacterial communities to an increase of K and Mg availability. Sci Rep 9:14403. https://doi.org/10.1038/s41598-019-50730-y.

Ontiveros-Cuadras, JF, Ruiz-Fernandez AC, Sanchez-Cabeza JA, Sericano J, Perez-Bernal LH, Paez-Osuna F, Dunbar RB, Mucciarone DA (2019) Recent history of persistent organic pollutants (PAHs, PCBs, PBDEs) in sediments from a large tropical lake. J Hazardous Mat 368, 264-273. https://doi.org/10.1016/j.jhazmat.2018.11.010.

Pagnout C, Frache G, Poupin P, Maunit B, Muller JF, Ferard JF (2007) Isolation and characterization of a gene cluster involved in PAH degradation in *Mycobacterium* sp. strain SNP11:expression in *Mycobacterium smegmatis* mc(2) 155. Res Microbiol 158:175–86

Parks DH, Beiko RG (2010) Identifying biologically relevant differences between metagenomic communities. Bioinformatics 26 (6):715-21

Patel KS, Ramteke S, Sahu BL, Nayak Y, Sharma S, Hung CC (2015) Polychlorinated Biphenyls Contamination of Sludge in India. AJAC 6:867-877. https://dx.doi.org/10.4236/ajac.2015.611082.

Randle-Boggis RJ, Helgason T, Sapp M, Ashton PD (2016) Evaluating techniques for metagenome annotation using simulated sequence data. FEMS Microbiol Ecol 92:95. https://doi.org/ 10.1093/femsec/fiw095.

Ribeiro H, Sousa T, Santos JP, Sousa AGG, Teixeira C, Monteiro MR, Salgado P, Mucha AP, Almeida CMR, Torgo L, Magalhaes C (2018) Potential of dissimilatory nitrate reduction pathways in polycyclic aromatic hydrocarbon degradation. Chemosphere 199:54-67

Rodrigues JL, Kachel CA, Aiello MR, Quensen JF, Maltseva OV, Tsoi TV, Tiedje JM (2006) Degradation of aroclor 1242 dechlorination products in sediments by *Burkholderia xenovorans* LB400 (ohb) and *Rhodococcus* sp. strain RHA1(fcb). Appl Environ Microbiol 72 (4):2476-82

Saavedra JM, Acevedo F, Gonzalez M, Seeger M (2010) Mineralization of PCBs by the genetically modified strain *Cupriavidus necator* JMS34 and its application for bioremediation of PCBs in soil. Appl Microbiol Biotechnol 87:1543–1554. https://doi.org/ 10.1007/s00253-010-2575-6.

Saito A, Iwabuchi T, Harayama S (2000) A novel phenanthrene dioxygenase from *Nocardioides* sp. strain KP7: expression in *Escherichia coli*. J Bacteriol 182:2134–2141

Samson R, Shah M, Yadav R, Sarode P, Rajput V, Dastager SG, Dharne MS, Khairnar K (2019) Metagenomic insights to understand transient influence of Yamuna River on taxonomic and functional aspects of bacterial and archaeal communities of River Ganges. J Sci Tot Env 674:288-299. https://doi.org/10.1016/j.scitotenv.2019.04.166.

Seo JS, Keum YS, Li QX (2009) Bacterial degradation of aromatic compounds. Int J Environ Res Public Health 6 (1):278–309. https://doi.org/10.3390/ijerph6010278.

Singh DP, Prabha R, Gupta VK, Verma MK (2018) Metatranscriptome analysis deciphers multifunctional genes and enzymes linked with the degradation of aromatic compounds and pesticides in the wheat rhizosphere. Front Microbiol 9:1331. https://doi.org/10.3389/fmicb.2018.01331.

Stanford RA, Cole JR, Tiedje JM (2002) Characterization and Description of *Anaeromyxobacter dehalogenans* gen. nov., sp. nov., an Aryl-Halorespiring Facultative Anaerobic *Myxobacterium.* Appl Environ Microbiol 68 (2):893-900. https://doi.org/ 10.1128/AEM.68.2.893-900.2002.

Su XM, Liu YD, Hashmi MZ, Ding LX, Shen CF (2015) Culture-dependent and culture-independent characterization of potentially functional biphenyl-degrading bacterial community in response to extracellular organic matter from Micrococcus luteus. Microb Biotechnol 8 (3): 569–578. https://doi.org/10.1111/1751-7915.12266

Valderrama JA, Durante-Rodriguez G, Blazquez B, Garcia JL, Carmona M, Díaz E (2012) Bacterial Degradation of Benzoate cross-regulation between aerobic and anaerobic pathways. J Biol Chem 287 (13):10494 –10508

Van Gerven T, Geysen D, Vandecasteele C (2004) Estimation of the contribution of a municipal waste incinerator to the overall emission and human intake of PCBs in Wilrijk,Flanders.Chemosphere 54:1303–1308.https://doi.org/10.1016/S0045-6535(03)00233-9.

Vila J, Nieto M, Mertens J, Springael D, Grifoll M (2010) Microbial community structure of a heavy fuel oil-degrading marine consortium, linking microbial dynamics with polycyclic aromatic hydrocarbon utilization. FEMS Microbiol Ecol 73:349–362

Vinas M, Sabate J, Espuny MJ, Solanas AM (2005) Bacterial Community Dynamics and Polycyclic Aromatic Hydrocarbon Degradation during Bioremediation of Heavily Creosote-Contaminated Soil. Appl Environ Microbiol 7008–7018. https://doi.org/ 10.1128/AEM.71.11.7008-7018.2005.

Weiland-Brauer N, Fischer MA, Schramm KW, Schmitz RA (2017) Polychlorinated Biphenyl (PCB)-Degrading Potential of Microbes Present in a Cryoconite of Jamtalferner Glacier. Front Microbiol 8:1105. https://doi.org/10.3389/fmicb.2017.01105.

Wittich RM, Strompl C, Moore ERB, Blasco R, Timmis KN (1999) Interactions of *Sphingomonas* and *Pseudomonas* strains in the degradation of chlorinated dibenzofurans. J Ind Microbiol Biotechnol 23:353-358

Yang HH, Lai SO, Hsieh LT, Hsueh HJ, Chi TW (2002) Profiles of PAH emission from steel and iron industries. Chemosphere 48: 1061–1074

Yang X, Liu X, Song L, Xie F, Zhang G, Qian S (2007) Characterization and functional analysis of a novel gene cluster involved in biphenyl degradation in *Rhodococcus* sp. strain R04. J Appl Microbiol 103:2214–2224. https://doi.org/ 10.1111/j.1365-2672.2007.03461.x

Zhang De-C, Mortelmaier C, Margesin R (2012) Characterization of the bacterial archaeal diversity in hydrocarbon-contaminated soil. Sci Tot Environ 421–422:184-196. https://doi.org/10.1016/j.scitotenv.2012.01.043.

Zhang W, Wei C, An G (2015) Distribution, partition and removal of polycyclic aromatic hydrocarbons (PAHs) during coking wastewater treatment processes. Environ Sci Process Impacts 17 (5):975-984

Zhang S, Hu Z, Wang H (2019) Metagenomic analysis exhibited the co-metabolism of polycyclic aromatic hydrocarbons by bacterial community from estuarine sediment. Environ Int 129:308-319

Tables

Table 1 Physio chemical parameters of the contaminated soil sample MGB-2 and MGB-3 from polluted near Bhilai steel plant

Table 2 Alpha diversity for bacterial, archaea and eukaryota communities of collected soil from polluted site near

Bhilai steel plant

Soil testing parameter	MGB-2	MGB-3	Reference value
pH	7.83	7.69	V. acidic < 5.0, Acidic 5.0 < 6.0, Normal 6.0 - 8.0, Alkaline 8.0 < 9.0
Electrical Conductivity (dSm-1)	0.21	0.183	<1.0 Normal
Organic Carbon (%)	0.85	1.39	< 0.5 Low, 0.50 - 0.75 Medium, > 0.75 - High
Nitrogen (kg ha-1)	430	734	< 280 Low, 281 - 560 Medium, > 560 High
Phosphorous(kg ha-1)	17.66	56.9	< 22 Low, 23 - 56 Medium, > 56 High
Potassium(kg ha-1)	1232	336	< 112 Low, 113 - 280 Medium, > 280 High
Calcium Carbonate (%)	7.2	7.6	< 1 Low, 1 - 5 Normal, 5-10 Sufficient, > 10 Harmful
Available Calcium (ppm)	640	960	< 500 Low, 500 -1000 Normal, >1000 Sufficient
Magnesium (ppm)	768	576	< 250 Low, 250 - 500 Normal, >500 Sufficient
Available Sodium (ppm)	138	126.5	Up to 400 Normal, 400 - 700 Problem may occur, > 700 Harmful
Chloride (ppm)	11.92	7.95	Up to 350 Normal, 350 - 1050 Slightly problem, > 1050 Harmful
Sulphur (mgkg-1)	16.37	25.37	< 10 Low, 10 -50 Normal, 50 High
Zinc (mgkg-1)	1.218	3.267	< 0.6 Low, 0.61- 5.0Medium, >5.1 High
Iron (mgkg-1)	8.372	14.98	< 4.5 Low, 4.6 -24 Medium, >25 High
Copper (mgkg-1)	1.286	1.575	< 0.2 Low, 0.3 -1.5 Medium, >1.5 High
Manganese (mgkg-1)	44.12	17.44	< 2.0 Low, 2.1 - 29 Medium, >30 High
Water Holding Capacity (%)	46	43.81	< 20 Low, 20 -50 Medium, > 50 High

Domain	Bacteria		Archaea		Eukaryota		Total Diversity	
Sample	MGB-2	MGB-3	MGB-2	MGB-3	MGB-2	MGB-3	MGB-2	MGB-3
Simpson_1-D	0.9957	0.9964	0.9788	0.9798	0.9847	0.987	0.9959	0.9966
Shannon_H	6.175	6.25	4.12	4.139	4.638	4.743	6.267	6.341
Evenness_e^H/S	0.3456	0.3689	0.7075	0.7298	0.4921	0.5338	0.3057	0.3232
Brillouin	6.155	6.23	4.03	4.034	4.492	4.614	6.244	6.318
Menhinick	3.247	3.168	1.849	2.024	4.17	3.856	3.972	3.91
Margalef	114.7	115.1	11.17	11.34	26.66	26.61	141.9	143.6
Equitability_J	0.8532	0.8624	0.9225	0.9293	0.8674	0.8831	0.841	0.8488
Fisher_alpha	204.5	204.4	18.06	18.8	54.35	52.45	262	264.4
Berger-Parker	0.02676	0.0224	0.05917	0.04873	0.04574	0.04471	0.02608	0.02184
Chao-1	1414	1429	87	86.33	266	267.1	1839	1884

Table 3 Annotated gene and enzyme identified in metagenome MGB-2 and MGB-3

a. methane metabolism

Metabolism	sublevel 3	gene	Enzyme identified
Methane	coenzyme M biosynthesis	*com*E	Sulfopyruvate decarboxylase - beta subunit [EC 4.1.1.79]
	coenzyme M biosynthesis	*com*A	Phosphosulfolactate synthase [EC 4.4.1.19]
	Hydrogenases	*hox*F	NAD-reducing hydrogenase subunit HoxF [EC 1.12.1.2]
	Hydrogenases	*hox*Y	NAD-reducing hydrogenase subunit HoxY [EC 1.12.1.2]
	CO Dehydrogenase	*cut*L	Carbon monoxide dehydrogenase form I, large chain [EC 1.2.99.2]
	H2:CoM-S-S-HTP oxidoreductase	*hdr*A	CoB–CoM heterodisulfide reductase subunit A [EC 1.8.98.1]
	CO Dehydrogenase	*cox*S	Carbon monoxide dehydrogenase small chain [EC 1.2.99.2]
	Hydrogenases	*hox*h	NAD-reducing hydrogenase subunit HoxH [EC 1.12.1.2]
	CBSS-314269.3.peg.1840	*cox*L	Carbon monoxide dehydrogenase large chain [EC 1.2.99.2]
	CBSS-314269.3.peg.1840	*cox*M	Carbon monoxide dehydrogenase medium chain [EC 1.2.99.2]
	Methanogenesis from methylated	*mtt*B	Trimethylamine:corrinoid methyltransferase [EC 2.1.1.250]
	Methanogenesis	*hdr*C2	CoB–CoM heterodisulfide reductase subunit C [EC 1.8.98.1]
	Serine-glyoxylate cycle	*mc*H	N(5),N(10)-methenyltetrahydromethanopterin cyclohydrolase [EC 3.5.4.27]
	Methanogenesis	*fhc*D	Formylmethanofuran--tetrahydromethanopterin N-formyltransferase [EC 2.3.1.101]
	One-carbon by tetrahydropterines	*mtd*C	Methylene tetrahydromethanopterin dehydrogenase [EC 1.5.99.9]
	Methanogenesis	*fwd*A	Formylmethanofuran dehydrogenase subunit A [EC 1.2.99.5]
	Methanogenesis	*fno*	F420-dependent N(5),N(10)-methylenetetrahydromethanopterin reductase [EC 1.5.99.11]
	Methanogenesis	*fwd*B	Formylmethanofuran dehydrogenase subunit B [EC 1.2.99.5]

b-c. Nitrogen and Sulfur metabolism

Table 4. Annotated enzyme and the assigned genera in metagenome MGB-2 and MGB-3

a. Chlorocyclohexane and chlorobenzene degradation pathway

Metabolism	sublevel 3	gene	Enzyme identified
Nitrogen	Denitrification	nirS	Cytochrome cd1 nitrite reductase [EC 1.7.2.1]
	Denitrification	nirK	Copper-containing nitrite reductase [EC 1.7.2.1]
	Denitrification	norB	Nitric-oxide reductase subunit B [EC 1.7.99.7]
	Denitrification	norC	Nitric-oxide reductase subunit C [EC 1.7.99.7]
	Denitrification	nosZ	Nitrous-oxide reductase [EC 1.7.99.6]
	Denitrification	norD	Nitric oxide reductase activation protein NorD
	Denitrification	norQ	Nitric oxide reductase activation protein NorQ
	Denitrification	qnor	Nitric-oxide reductase, quinol-dependent [EC 1.7.99.7]
	Nitrate and nitrite ammonification	nitH	Polyferredoxin NapH (periplasmic nitrate reductase)
	Nitrate and nitrite ammonification	nirB	Assimilatory nitrate reductase large subunit [EC:1.7.99.4]
	Nitrate and nitrite ammonification	nitH	Nitrite reductase [NAD(P)H] small subunit [EC 1.7.1.4]
	Nitrate and nitrite ammonification	nosZ	Ferredoxin--nitrite reductase [EC 1.7.7.1]
	Nitrate and nitrite ammonification	*nrfE*	Cytochrome c-type heme lyase subunit nrfE [EC 4.4.1.-]
	Nitrate and nitrite ammonification	napB	Nitrate reductase cytochrome c550-type subunit [EC 1.9.6.1]
	Nitrate and nitrite ammonification	nitH	Nitrite reductase [NAD(P)H] large subunit [EC 1.7.1.4]
	Nitrate and nitrite ammonification	napA	Periplasmic nitrate reductase precursor [EC 1.7.99.4]
	Nitrogen fixation	nifN	Nitrogenase FeMo-cofactor scaffold and assembly protein NifN [EC 1.18.6.1]
	Nitrogen fixation	nifE	Nitrogenase FeMo-cofactor scaffold and assembly protein NifE[EC 1.18.6.1]
	Nitrogen fixation	nifK	Nitrogenase (molybdenum-iron) beta chain [EC 1.18.6.1]
Sulfur	Inorganic Sulfur Assimilation	cysN	Sulfate adenylyltransferase, dissimilatory-type [EC 2.7.7.4]
	Inorganic Sulfur Assimilation	cysH	Phosphoadenylyl-sulfate reductase [thioredoxin] [EC 1.8.4.8]
	Inorganic Sulfur Assimilation	siR	Ferredoxin--sulfite reductase [EC 1.8.7.1]
	Inorganic Sulfur Assimilation	papsR	Adenylyl-sulfate reductase [thioredoxin] [EC 1.8.4.10]
	Inorganic Sulfur Assimilation	cysC	Adenylylsulfate kinase [EC 2.7.1.25]
	Inorganic Sulfur Assimilation	cysI	Sulfite reductase [NADPH] hemoprotein beta-component [EC 1.8.1.2]

Pathway	Gene	Anotated Enzyme	Assigned bacteria in MGB-2	Assigned bacteria in MGB-3
Chlorocyclohexane and chlorobenzene degradation [PATH:ko00361]	chqB	hydroxyquinol 1,2-dioxygenase [EC:1.13.11.37]	*Arthrobacter, Bradyrhizobium Burkholderia, Cupriavidus Variovorax , Verminephrobacter*	*Catenulispora Bradyrhizobium*
	pcpB	pentachlorophenol monooxygenase [EC:1.14.13.50]	*Rhodococcus*	*Anabaena, Burkholderia*
	pheA	phenol 2-monooxygenase [EC:1.14.13.7]	*Arthrobacter, Bradyrhizobium, Nocardia, Renibacterium, Rhodococcus, Rubrobacter*	*Agrobacterium, Arthrobacter, Azoarcus, Bradyrhizobium, Methylobacterium, Nocardia, Paracoccus, Rhodococcus, Rubrobacter*
	pnpD	maleylacetate reductase [EC:1.3.1.32]	*Bradyrhizobium, Polaromonas*	*Bradyrhizobium, Burkholderia, Polaromonas, Rhodococcus, Rhodopseudomonas*
	clcD	carboxymethylenebutenolidase [EC:3.1.1.45]	*Acidimicrobium , Agrobacterium, Arthrobacter, Azoarcus, Bradyrhizobium, Burkholderia, Candidatus, Cupriavidus, Desulfobacterium, Gemmatimonas, Methanoculleus, Polaromonas, Pseudomonas, Rhodococcus, Rhodopseudomonas, Sphingobium*	*Acetobacter, Acidimicrobium, Acidobacterium, Agrobacterium, Azoarcus, Beijerinckia, Bradyrhizobium, Burkholderia, Candidatus, Cellvibrio, Cupriavidus, Dechloromonas, Desulfobacterium,, Flavobacterium, Gemmatimonas, Polaromonas, Rhodococcus, Sphingomonas*
	hadL	2-haloacid dehalogenase [EC:3.8.1.2]	*Beijerinckia, Bordetella, Burkholderia, Cupriavidus, Methylobacterium, Mycobacterium, Oligotropha, Rhodopseudomonas*	*Acidovorax, Anaeromyxobacter, Beijerinckia, Bordetella, Bradyrhizobium, Burkholderia, Candidatus, Chromobacterium, Mesorhizobium, Methylobacterium, Mycobacterium, Methylocella, Polaromonas, Rhodopseudomonas, Sorangium, Xanthomonas*
	dehH	haloacetate dehalogenase [EC:3.8.1.3]	*Anabaena, Anaeromyxobacter, Azoarcus, Burkholderia, Chloroflexus, Cupriavidus, Kribbella, Mesorhizobium, Methylobacterium, Nocardiopsis, Micromonospora Psychromonas, Rhizobium, Rhodopseudomonas, Roseiflexus, Saccharopolyspora*	*Anabaena, Anaeromyxobacter, Bradyrhizobium, Burkholderia, Chloroflexus, Geodermatophilus,, Mesorhizobium, Methylibium, Methylobacterium, Nocardiopsis, Psychromonas, Rhodopseudomonas, Roseiflexus,*
	bedC1 todC1	benzene/toluene dioxygenase subunit alpha [EC:1.14.12.3 1.14.12.11]	*Mycobacterium*	*Mycobacterium*
	bphC	biphenyl-2,3-diol 1,2-dioxygenase [EC:1.13.11.39]	*Rhizobium*	*Mycobacterium, Candidatus, Bradyrhizobium*
	catA	catechol 1,2-dioxygenase [EC:1.13.11.1]	*Cupriavidus, Methylobacterium, Pseudomonas, Sphingobium, Streptomyces*	*Azospirillum, Bradyrhizobium, Burkholderia, Delftia, Mycobacterium, Pseudomonas, Rhizobium, Sphingomonas,*
	catB	muconate cycloisomerase [EC:5.5.1.1]	*Rhodopirellula*	*Mycobacterium*
	dmpB	catechol 2,3-dioxygenase [EC:1.13.11.2]	*Arthrobacter, Bradyrhizobium, Brucella, Geobacillus, Meiothermus, Rubrobacter*	*Agrobacterium, Bradyrhizobium, Burkholderia, Meiothermus, Methylibium, Rhodococcus,*

b. Biphenyl/ PAH degrading pathway

Pathway	Gene	Annotated Enzymes	Assigned bacteria in MBG-2	Assigned bacteria in MBG-3
Biphenyl degradation	*bph*A	biphenyl 2,3-dioxygenase [EC:1.14.12.18]	*Mycobacterium vanbaalenii* PYR-1	*Mycobacterium vanbaalenii* PYR-1, *Polaromonas naphthalenivorans* CJ2
	*bph*B	cis-2,3-dihydrobiphenyl-2,3-diol dehydrogenase [EC 1.3.1.56]	*nd*	*nd*
	*bph*C	biphenyl-2,3-diol 1,2-dioxygenase [EC:1.13.11.39]	*Mycobacterium smegmatis* str. MC2 155	*Alkalilimnicola ehrlichii* MLHE-1, Mycobacterium smegmatis str. MC2 155, *Mycobacterium tuberculosis* KZN 1435, *Mycobacterium bovis* BCG str. Tokyo 172
	*bph*D	2,6-dioxo-6-phenylhexa-3-enoate hydrolase [EC:3.7.1.8]	*Nocardioides* sp. JS614	*Mycobacterium marinum* M, *Polynucleobacter necessaries,*
	*bph*E	2-hydroxypenta-2,4-dienoate hydratase [EC 4.2.1.80]	*nd*	*Polynucleobacter necessarius*
	*bph*F	4-hydroxy-2-oxovalerate aldolase [EC 4.1.3.39]	*Carboxydothermus hydrogenoformans, Pseudomonas putida* F1, *Burkholderia* sp. 383	*Carboxydothermus hydrogenoformans* Z-2901, *Legionella pneumophila* str. Corby *Mycobacterium* sp. JLS *Roseiflexus* sp. RS-1 *Arcobacter butzleri* *Rhodococcus jostii* RHA1 *Streptosporangium roseum* DSM 43021 *Nocardioides* sp. JS614 *Nocardia farcinica* IFM 10152 *Eubacterium rectale* ATCC 33656 *Frankia* sp. EAN1pec
PAH degradation	*nahD*	hydroxychromene-2-carboxylate isomerase [EC:5.99.1.4]	*Agrobacterium tumefaciens,* *Aromatoleum aromaticum, Bradyrhizobium sp.* BTAi1, *Dechloromonas aromatic, Pseudoalteromonas atlantica*	*Aromatoleum aromaticum,* *Bradyrhizobium sp.* BTAi, *Burkholderia xenovorans, Parvibaculum lavamentivorans, Rhizobium leguminosarum, Ruegeria pomeroyi* *nd*
	nahAc	naphthalene dioxygenase ferredoxin [EC:1.14.12.12]	*Mesorhizobium loti*	*Acidovorax citrulli, Bordetella bronchiseptica, Corynebacterium efficiens, Cupriavidus pinatubonensis, Delftia acidovorans, Polaromonas sp. JS666, Streptomyces coelicolor*
	nahG	salicylate hydroxylase [EC:1.14.13.172]	*Anaeromyxobacter dehalogenans, Burkholderia cenocepacia, Dyadobacter fermentans Methylobacterium extorquens, Mycobacterium smegmatis, Polaromonas sp.* JS666, *Pseudoalteromonas atlantica*	*Mycobacterium* sp. JLS *Mycobacterium* sp. JLS *Mycobacterium vanbaalenii*
		PAH dioxygenase large subunit [EC:1.13.11.-]	nd *nd*	*Mycobacterium vanbaalenii*
	nidA	PAH dioxygenase small subunit [EC:1.13.11.-]	*Mycobacterium vanbaalenii*	*Mycobacterium* sp. KMS
	nidB	aldehyde dehydrogenase [EC:1.2.1.-] extradiol dioxygenase [EC:1.13.11.-]	*nd*	*Mycobacterium vanbaalenii*
	nidD	hydratase-aldolase [EC:4.1.2.-]	*nd*	*Mycobacterium sp. MCS*
	phdF	1-hydroxy-2-naphthoate dioxygenase [EC:1.13.11.38]	*Mycobacterium vanbaalenii*	
	phdG	4-(2-carboxyphenyl)-2-oxobut-3-enoate aldolase [EC:4.1.2.34]	*Mycobacterium sp. KMS,* *Mycobacterium sp. MCS*	
	phdI			

	phdJ

c. benzoate degradation *via* benzoyl coA, protocatechuate and catechol pathways

Pathway	Gene	Annotated Enzymes	Assigned bacteria in MBG-2	Assigned bacteria in MBG-3
Benzoate degradation	*ben*B/ *xyl*Y	benzoate/toulene1,2 dioxygenase beta subunit	*Rhodopseudomonas,Rhodococcus*	*Geodermatophilus, Saccharo*
	*ben*C/ *xyl*Z	benzoate/toulene 1,2 dioxygenase electron transfer component	*nd*	*Geodermatophilus, Mycobac Saccharopolyspora*
	*ben*D/ DHBD	2,3-dihydroxybenzoate decarboxylase [EC:4.1.1.46]	*Burkholderia,* *Rhodococcus, Rhodopseudomonas*	*Bordetella, Novosphingobiu Polaromonas, Rhizobium, Rhodopseudomonas, Xanth*
	*pim*C	pimeloyl-CoA dehydrogenase [EC:1.3.1.62]	*Cupriavidus*	*Cupriavidus*
	*pra*C	4-oxalocrotonate tautomerase [EC:5.3.2.-]	*nd*	*Candidatus, Methylocella*
	*ali*A	cyclohexanecarboxylate-CoA ligase [EC:6.2.1.-]	*nd*	*Conexibacter, Cupriavidus, Rhodopseudomonas, Vermi*
	*bad*A	benzoate-CoA ligase [EC:6.2.1.25]	*Achromobacter, Aromatoleum, Azoarcus, Burkholderia, Comamonas, Cupriavidus, Delftia*	*Albidiferax , Amycolatopsis, Azoarcus, Bradyrhizobium, C Leptothrix, Rhodomicrobium*
	*bad*D	benzoyl-CoA reductase subunit [EC:1.3.7.8]	*Rhodopseudomonas, Thauera*	*Rhodopseudomonas*
	*bad*E	benzoyl-CoA reductase subunit [EC:1.3.7.8]	*Magnetospirillum , Rhodomicrobium*	*nd*
	*bad*F	benzoyl-CoA reductase subunit [EC:1.3.7.8]	*Magnetospirillum , Rhodomicrobium , Rhodopseudomonas, Thauera*	*Magnetospirillum , Rhodomi*
	*bad*H	2-hydroxycyclohexanecarboxyl-CoA dehydrogenase [EC:1.1.1.-]	*Cupriavidus, Magnetospirillum, Nocardioides*	*Alicyclobacillus, Aromatoleu Rhodopseudomonas*
	*bad*I	2-ketocyclohexanecarboxyl-CoA hydrolase [EC:3.1.2.-]	*Nd*	*Cupriavidus, Leptothrix , Pol Xanthobacter*
	*hba*A	4-hydroxybenzoate-CoA ligase [EC:6.2.1.27 6.2.1.25]	*Bradyrhizobium , Magnetospirillum, Rhodopseudomonas*	*Bradyrhizobium*
	*hba*B, *hcr*C	4-hydroxybenzoyl-CoA reductase subunit gamma [EC:1.3.7.9]	*Magnetospirillum, Rhodopseudomonas*	*nd*
	*hba*C, *hcr*A	4-hydroxybenzoyl-CoA reductase subunit alpha [EC:1.3.7.9]	*Rhodopseudomonas, Thauera*	*nd*
	*lig*AB	protocatechuate 4,5-dioxygenase [EC 1.13.11.8]	*Rhodopseudomonas, Acidimicrobium, Burkholderia, Nakamurella, Novosphingobium, Pseudoalteromonas, Xanthomonas*	*Brevundimonas, Azoar Brevundimonas, Marinomo Saccharomonospora*
	*lig*C	2-hydroxy-4-carboxymuconate semialdehyde hemiacetal dehydrogenase [EC 1.1.1.312]	*Rhodopseudomonas*	*Brevundimonas*
	*lig*I	2-pyrone-4,6-dicarboxylate lactonase [EC:3.1.1.57]	*Agrobacterium, Albidiferax, Arthrobacter, Bradyrhizobium, Novosphingobium, Rhodopseudomonas*	*Agrobacterium, Arthrobacter, Bradyrhizobium, Delftia, Pol Rhodopseudomonas*
	*lig*J	4-oxalomesaconate hydratase [EC 4.2.1.83]	*Arthrobacter, Asticcacaulis, Bradyrhizobium, Brevundimonas, Burkholderia, Cupriavidus, Burkholderia, Methylobacterium, Paracoccus, Pseudoalteromonas, Ralstonia , Rhizobium, Rhodopseudomonas, Sphingobium, Variovorax, Xanthomonas*	*Azoarcus,Brevundimonas, B Magnetospirillum,Methyloba Novosphingobium, Paracoc Polaromonas, Rhizobium, RhodopseudomonasVermin Xanthomonas*
	pobA	p-hydroxybenzoate 3-monooxygenase [EC:1.14.13.2]	*Amycolatopsis, Arthrobacter, Bacillus, Burkholderia, Candidatus, Caulobacter, Chelativorans Delftia, Frankia, Herbaspirillum, Mesorhizobium, Novosphingobium, Rhodopseudomonas, Sphaerobacter, Streptosporangium*	*Agrobacterium, Amycolatop Arthrobacter, Burkholderia, C Magnetospirillum, Pseudom Rhodopseudomonas, Saccharomonospora, Xanth Verminephrobacter*
	*pca*G	beta-Carboxy-cis,cis-muconate [EC:1.13.11.3]	*Bradyrhizobium, Candidatus,Cupriavidus, Geodermatophilus, Sphaerobacter, Sphingomonas*	*Candidatus*
	*pca*H	protocatechuate 3,4-dioxygenase, beta subunit [EC:1.13.11.3]	*Acidiphilium, Amycolatopsis, Arthrobacter, Azospirillum, Burkholderia, Caulobacter, Klebsiella, Pirellula, Planctomyces, Polaromonas, Pseudomonas, Rhodopirellula, Rubrobacter, Saccharopolyspora, Spirosoma, Streptosporangium, Xanthobacter*	*Agrobacterium, Azorhizobiu Azospirillum, Burkholderia , Cupriavidus, Geodermatoph Klebsiella, Mesorhizobium, F Polaromonas, Pseudomo Rhodococcus , Rhodopirellu Saccharopolyspora, Spiroso Streptomyces, Streptosporar Xanthomonas,*

	pcaI	3-oxoadipate CoA-transferase, alpha subunit [EC:2.8.3.6]	*Achromobacter , Arthrobacter, Kocuria, Serratia, Shewanella, Xanthomonas*	*Catenulispora, Mycobacteriu*
	pcaJ	3-oxoadipate CoA-transferase, beta subunit [EC:2.8.3.6]	*Arthrobacter , Cupriavidus, Escherichia , Herbaspirillum*	*Nocardioides, Rhodopseudo Ruegeria, Streptomyces*
	pcaB	3-carboxy-cis,cis-muconate cycloisomerase [EC:5.5.1.2]	*Cupriavidus, Deinococcus, Delftia, Frankia, Pseudomonas, Sinorhizobium, Xanthomonas*	*Albidiferax, Burkholderia, Ch Cupriavidus, Leptothrix, Mei Methylobacterium, Polarom Ralstonia , Xanthomonas*
	pcaC	4-carboxymuconolactone decarboxylase [EC:4.1.1.44]	*Beutenbergia, Bradyrhizobium, Burkholderia, Candidatus, Methylobacterium, Mycobacterium, Rhizobium, Rhodomicrobium, Sinorhizobium, Yersinia*	*Amycolatopsis, Arthrobacter Bradyrhizobium, Burkholder Mycobacterium, Ralstonia, Rhodopseudomonas, Sinorh*
	pcaD	3-oxoadipate enol-lactonase [EC:3.1.1.24]	*Burkholderia, Rhizobium, Xanthobacter*	*Chelativorans, Cupriavidus, Geodermatophilus, Methylob Mycobacterium, Rhizobium Sinorhizobium, Thermomon*
	pcaF	3-oxoadipyl-CoA thiolase [EC:2.3.1.174]	*Bradyrhizobium, Pseudomonas, Rhodopseudomonas*	*Bradyrhizobium, Sphingomo*
	catA	catechol 1,2-dioxygenase [EC:1.13.11.1]	*Cupriavidus, Methylobacterium, Pseudomonas, Sphingobium, Streptomyces*	*Azospirillum,Bradyrhizobium Delftia,Mycobacterium, Pseu Rhizobium, Sphingomona, X*
	catB	muconate cycloisomerase [EC:5.5.1.1]	*Rhodopirellula*	*Mycobacterium*
	catE/ dmpB	catechol 2,3-dioxygenase [EC:1.13.11.2]	*Agrobacterium, Bradyrhizobium, Burkholderia, Meiothermus, Methylibium, Rhodococcus, Thauera*	*Arthrobacter,Bradyrhizobium Geobacillus, Meiothermus, R*

Figures

Figure 1

Map illustrating the sampling location of soil samples MGB-2 and MGB-3 collected from a polluted site near Bhilai steel plant, Chattisgarh (India). Note: The designations employed and the presentation of the material on this map do not imply the expression of any opinion whatsoever on the part of Research Square concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. This map has been provided by the authors.

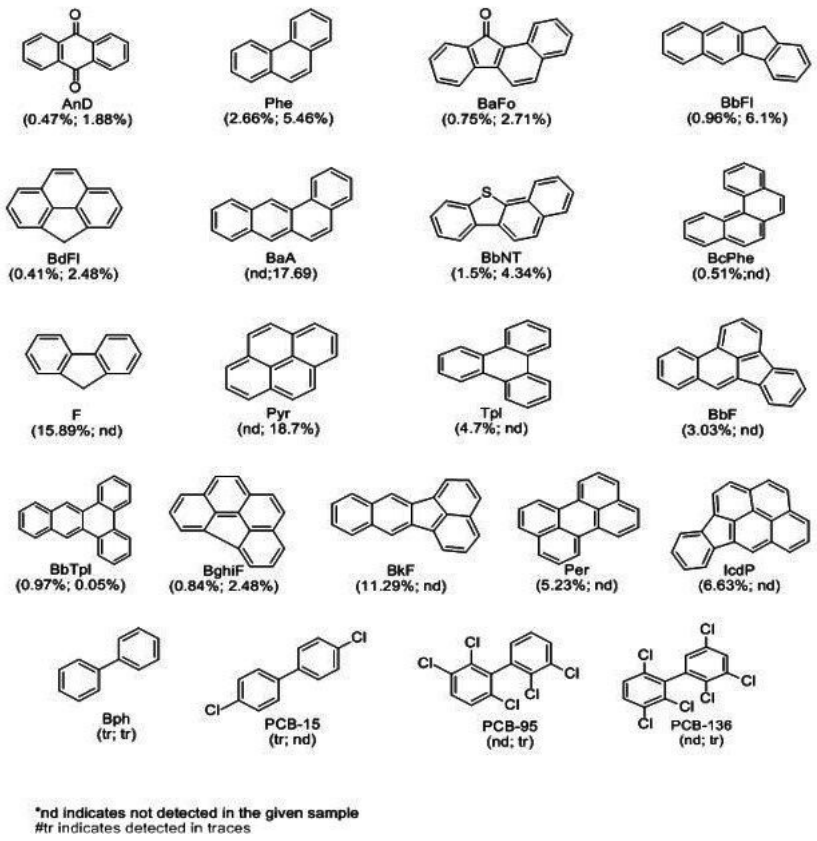

Figure 2

Structure and relative abundance of the PAH and PCB identified. PAH found in MGB-2 & MGB-3 were 9,10-Anthracenedione (AnD), 11H-Benzo[a]fluoren-11-one (BaFo), 11H-Benzo[b]fluorene (BbFl), Benzo[def]fluorene (BbFl), Benzo[b]fluoranthene (BbF), Benzo[b]naphtho[2,1-d]thiophene (BbNT), Benzo[b]triphenylene (BbTPl), Benzo[c]phenanthrene (BcPhe), Benzo[ghi]fluoranthene (BghiF), Benzo[k]fluoranthene (BkF), Fluorene (F), Indeno[1,2,3-cd] pyrene (IcdP), Perylene (Per), Phenanthrene (Phe), Pyrene (Pyr) and Triphenylene (Tpl) and are classified based on the numbers of aromatic rings 2, 3, 4, 5 and 6 membered rings. Bph, PCB-15, PCB-95 and PCB-136 represents biphenyl, 1,1'-biphenyl, 4,4'-dichloro, biphenyl, 1,1'-biphenyl 2,2',3,',5,6 pentachloro and 1,1'-biphenyl 2,2',3,3',6,6' hexachloro respectively.

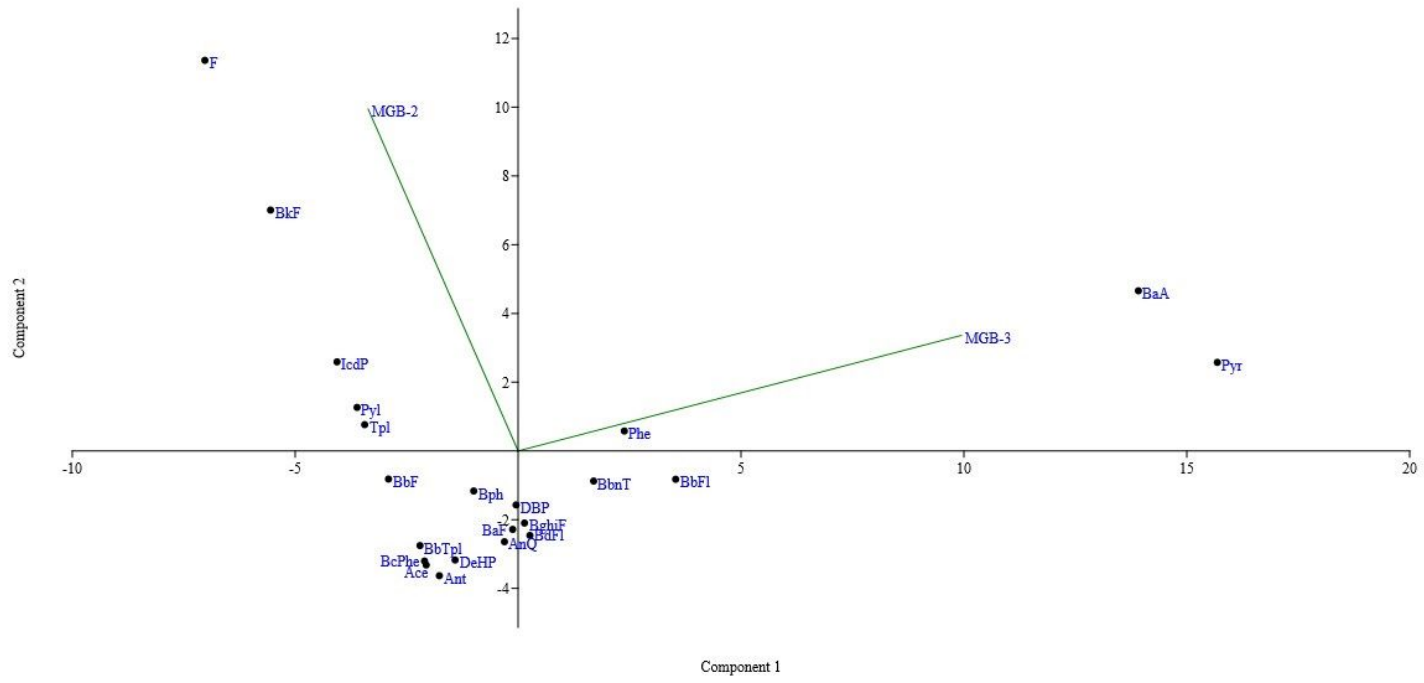

Figure 3

PCoA plot depicting relative abundance of PAH in MGB-2 and MGB-3 using Bray-Curtis matrix

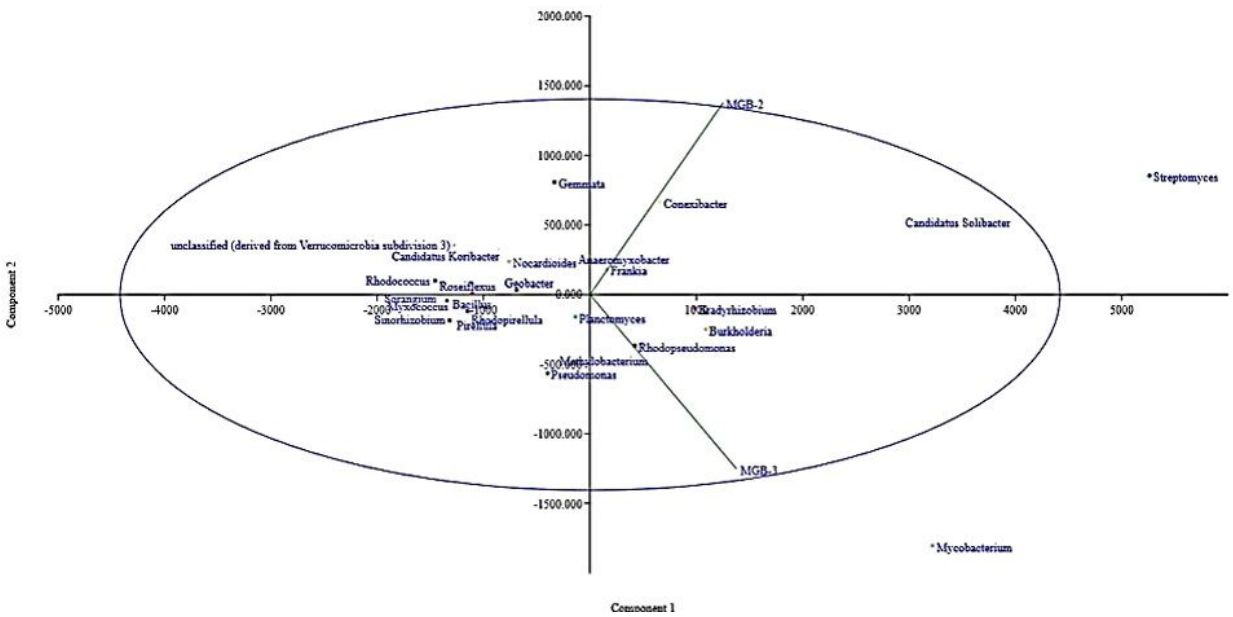

Figure 4

PCA plot depicting taxonomic analysis beta diversity of MGB-2 and MGB-3 on Bary -Curtis dissimilarity

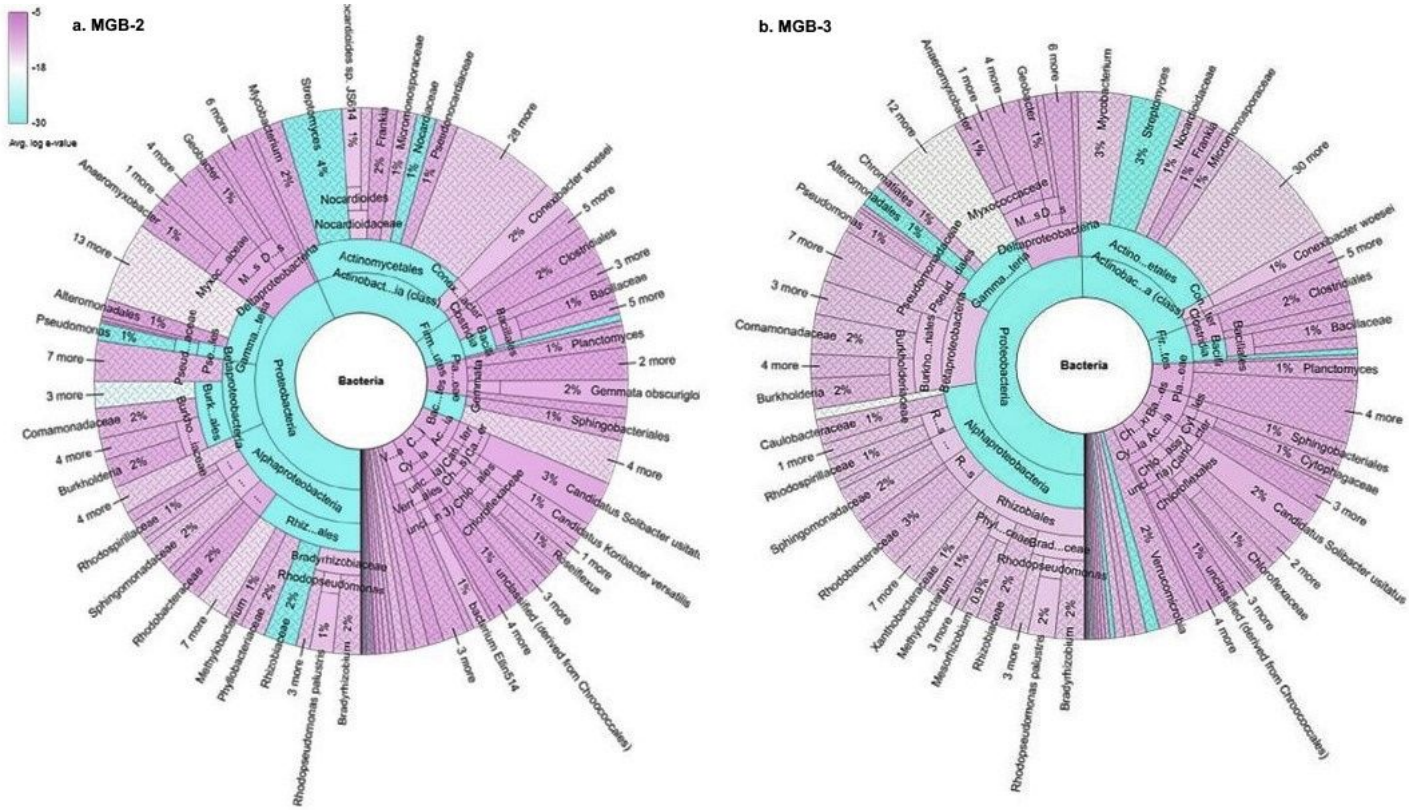

Figure 5

Krona plot demonstrated the relative abundance of taxa of bacteria across phylum to genus level hierarchy of a. MGB-2; b. MGB-3

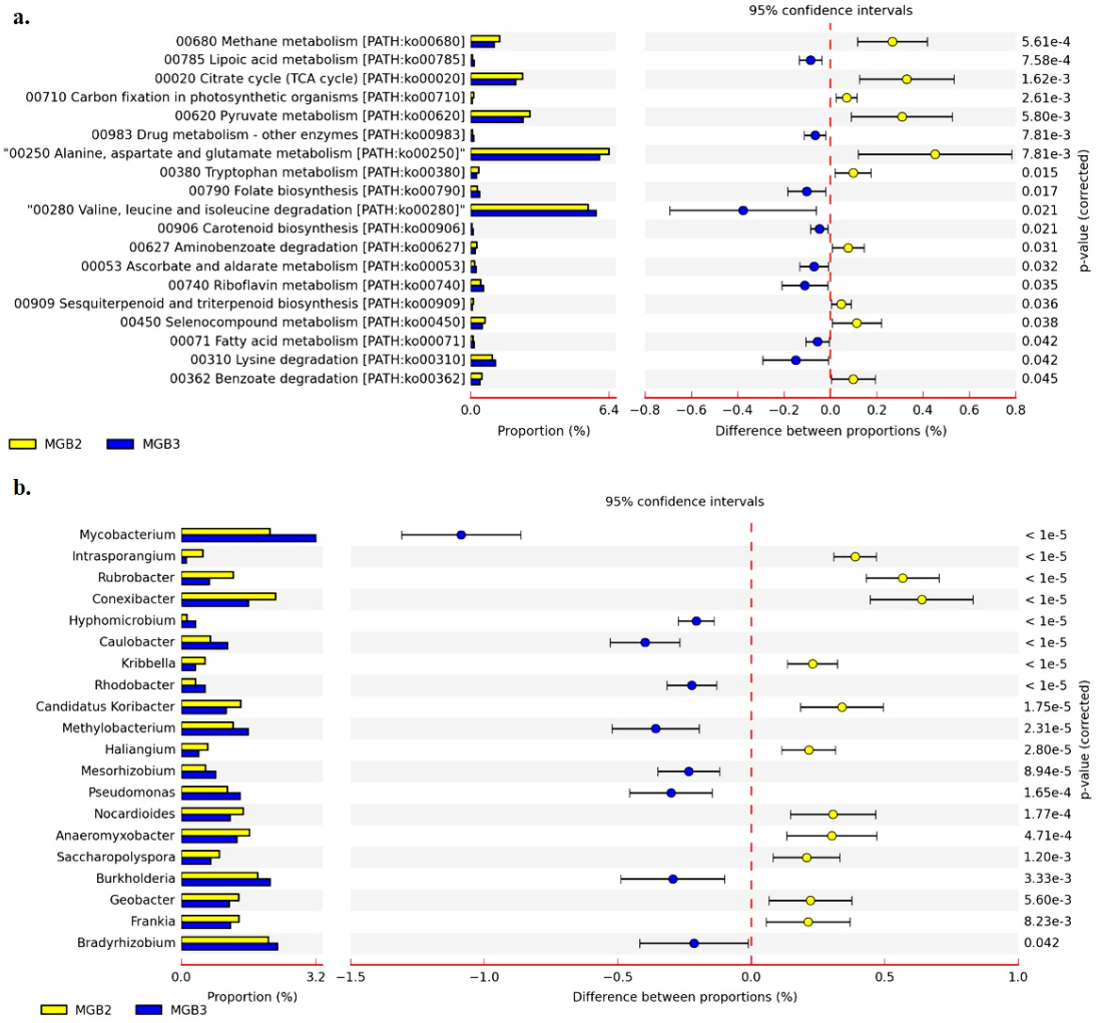

Figure 6

Comparative analysis of metagenome MGB-2 and MGB-3 using STAMP a.) SEED level based on metabolism b.) genus level with ReFSeq for metabolism

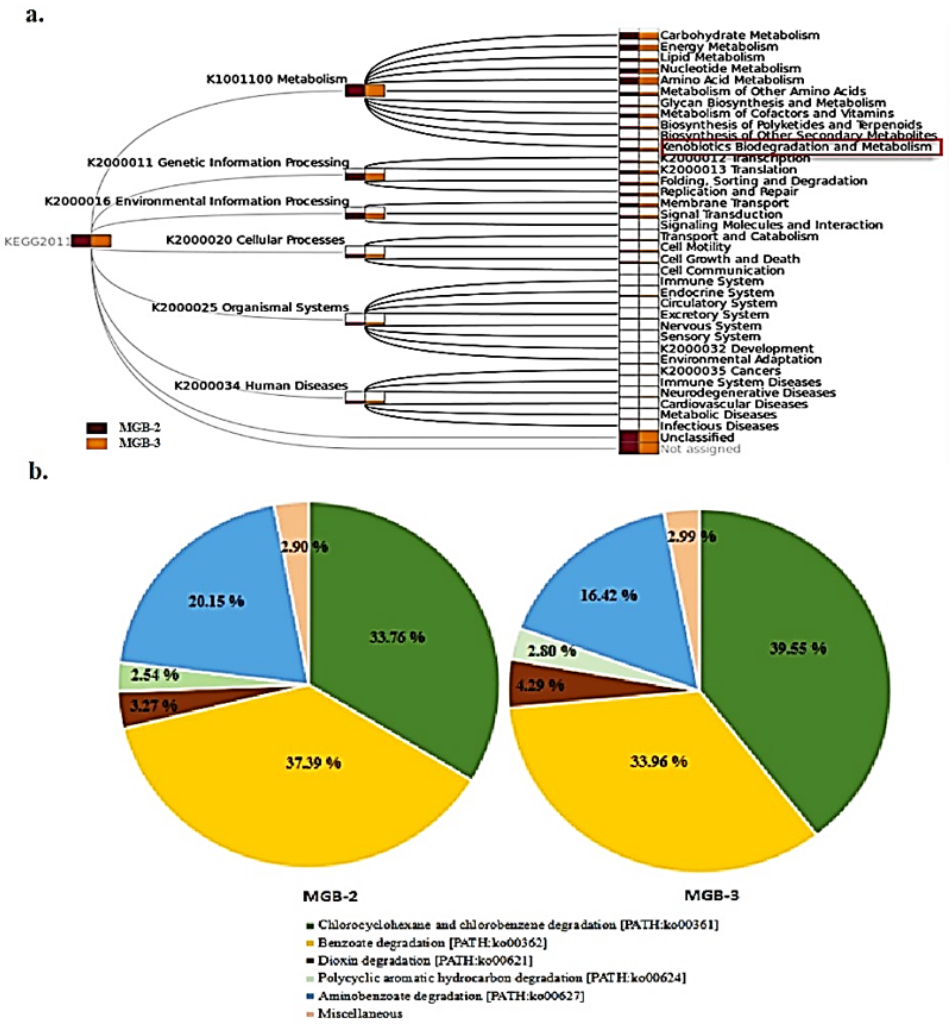

Figure 7

The SEED subsystem analysis in MG-RAST assigned reads in MGB-2 and MGB-3 based on various a.) functions b.) Xenobiotic biodegradation pathways

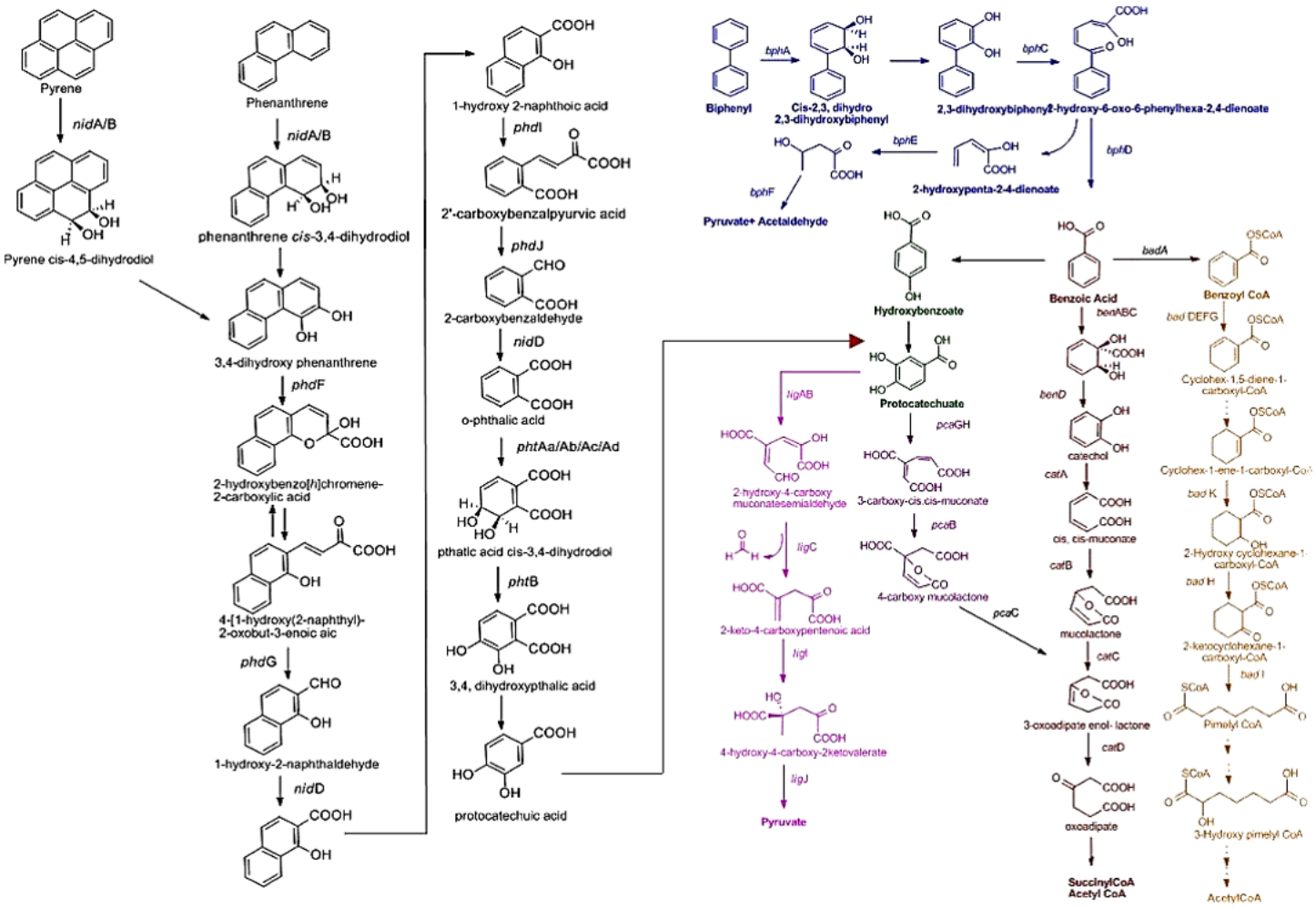

Figure 8

Reconstruction of complete Biphenyl/ PCB and PAH degradation pathways based on annotated genes identified. Blue, biphenyl degradation; red, benzoate degradation via catechol; pink, benzoate degradation via protocatechuate; orange, benzoate degradation via benzoyl CoA degradation; Black, PAH (Pyrene and Phenanthrene) degradation enter central pathway via protocatechauate intermediate

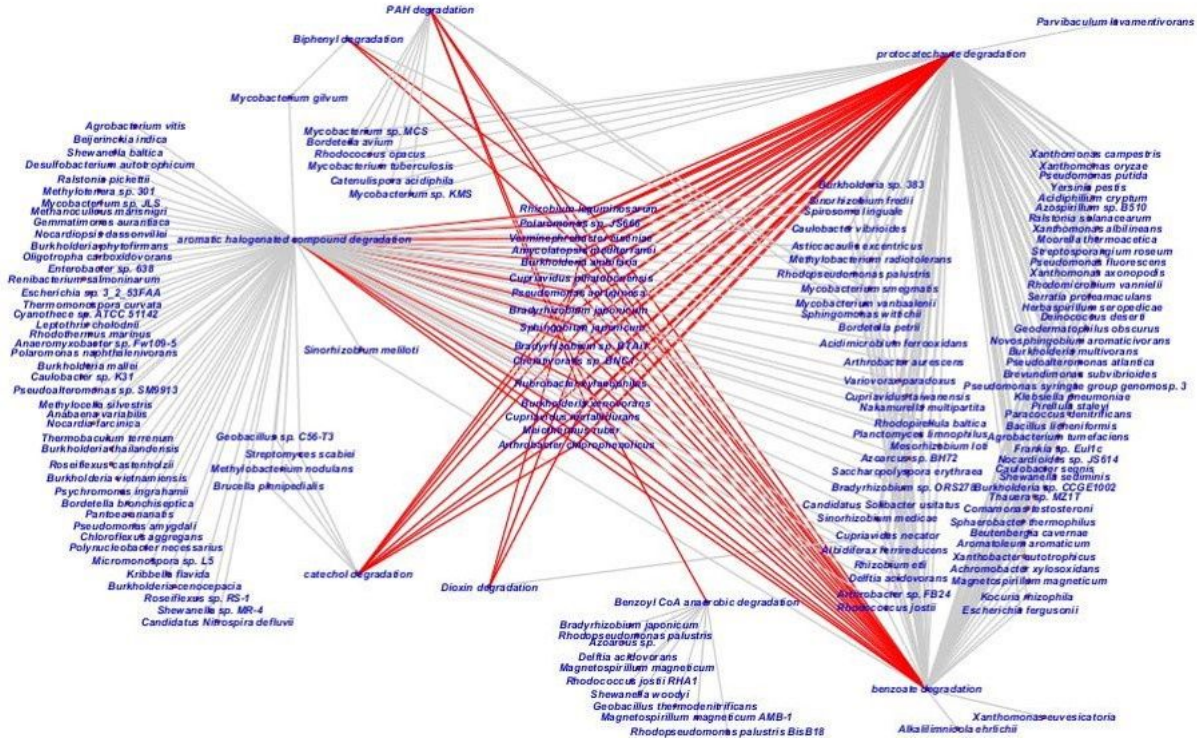

Figure 9

Cytoscape-based networking depicted interrelationship of key biodegrader in PAH, biphenyl, dioxin, halogenated, catechol and protocatechol pathways in MGB-2. Key biodegrader presented in the middle and the interrelated pathways are given in red lines

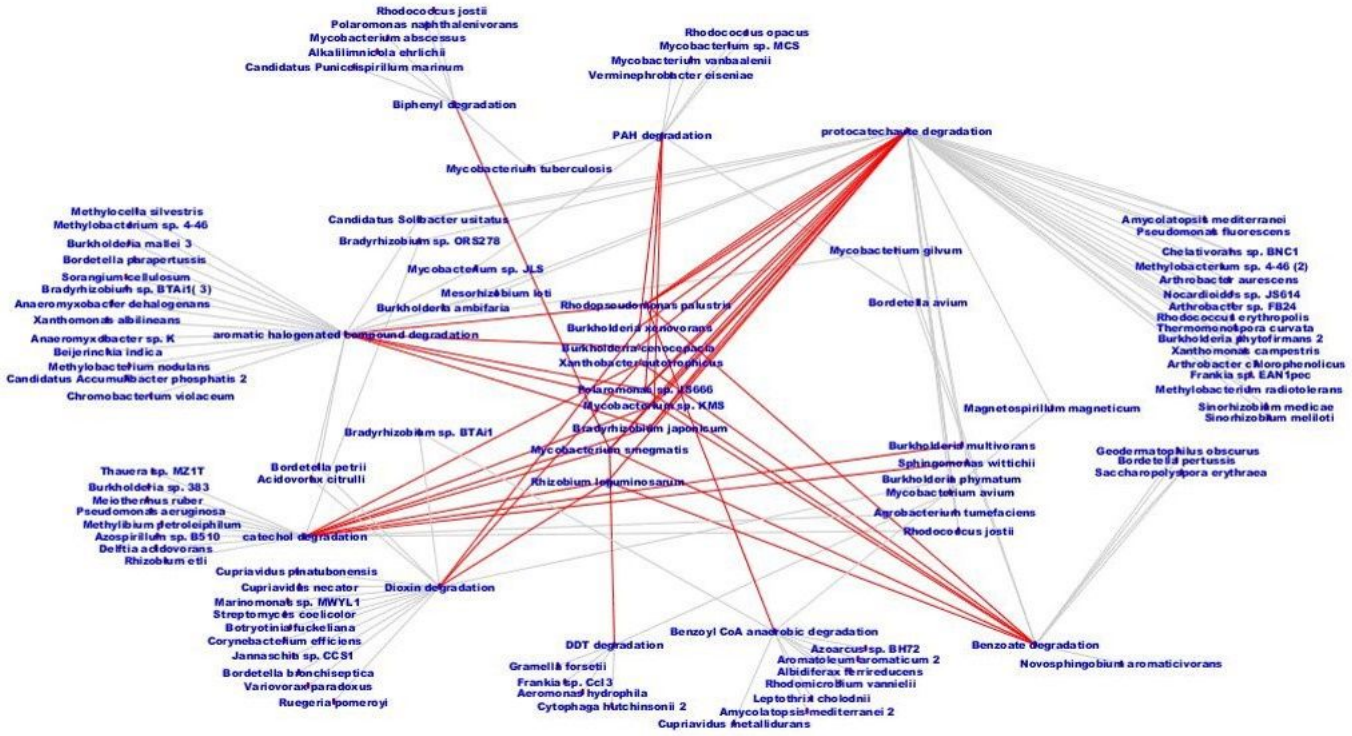

Figure 10

Cytoscape-based networking depicted interrelationship of key biodegrader in PAH, biphenyl, dioxin, halogenated, catechol and protocatechol pathways in MGB-3. Key biodegrader presented in the middle and the interrelated pathways are given in red lines

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- supplementaryData.docx