

SCIENTIFIC REPORTS



OPEN

Metagenomic profiling for assessing microbial diversity and microbial adaptation to degradation of hydrocarbons in two South African petroleum-contaminated water aquifers

Leonard Kachienga¹, Keshri Jitendra² & Maggy Momba¹

Biodegradation of hydrocarbons by indigenous populations of microorganisms found in petroleum-contaminated water sources represents one of the primary mechanisms by which petroleum and other hydrocarbon pollutants are eliminated from the aquatic environment. The identification of these microorganisms, which have capabilities to convert the majority of toxic hydrocarbons into compounds that are less harmful for end-users, is therefore crucial for bioremediation purposes. The aim of this study was to profile the microbial diversity of two South African petroleum-contaminated water aquifer sites and to determine the microbial adaptation to hydrocarbon degradation using a metagenomics approach. The sequenced samples revealed that protozoa (62.04%) were found to be the most dominant group, followed by fungi (24.49%), unknown (12.87%), and finally other sequences such as Animalia and plantae which were <(0.10%) domains in the first oil-polluted aquifer site. In the second site, protozoa (61.90%), unknown (16.51%), fungi (11.41%) in that order. According to the classification at the genus level, the dominant group was *Naegleria* (15.21%), followed by *Vorticella* (6.67%) as the only ciliated protozoan genus, other species such as *Arabidopsis* (2.97%), *Asarum* (1.84%) *Populus* (1.04%) were significantly low and drastically lower in the first site. Regarding the second site, the dominant group was *Naegleria* (18.29%) followed by *Colpoda* (9.86%) with the remainder of the genera representing <2%. Overall results demonstrated the ability of various groups of microorganisms to adapt and survive in petroleum oil-polluted water sites regardless of their respective distributions and this can be explored further for their role in bioremediation and environmental management.

Continuous pollution of the existing freshwater sites across the globe by hydrocarbons of petroleum and their by-products has disastrous consequences for the end-users such as animals, humans, plants and microorganisms. According to Haritash and Kaushik¹, the ongoing pollution by these hydrocarbons and their derivatives results in the accumulation of xenobiotics and other toxic compounds in both water sources and the entire environmental landscape. Petroleum is a complex mixture of hydrocarbons and non-hydrocarbons from metallic porphyrins, acid and organometallic compounds origin². Polycyclic aromatic hydrocarbons (PAHs) are among the toxic components of petroleum oil spills considered to be the main contributor of mutagenicity and teratogenicity in humans¹. Widespread oil spill pollution in various water sources has been reported during the 20th century; oil pollution has been found to be mainly of anthropogenic origin, such as leaks and spills due to oil refining, handling, storage and transport from the refineries to the point of usage³. Millions of tons of petroleum oil are being exploited as a source of energy for the ever-increasing energy demand by the rapidly growing world population⁴.

¹Department of Environmental, Water and Earth Sciences, Faculty of Science, Tshwane University of Technology, Arcadia Campus, P/Bag X680, Pretoria, 0001, South Africa. ²Institute of Postharvest and Food Sciences, Agricultural Research Organization, Volcani Centre, 50250, Bet Dagan, Israel. Correspondence and requests for materials should be addressed to M.M. (email: mombamnb@tut.ac.za)

Microbial diversity involved in bioremediation is always considered as a form of natural degradation of hydrocarbons in various water sources³. The use of bioremediation techniques, which are mainly based on the use of microorganisms to biodegrade the crude oil contaminants, has shown signs of being a promising technology due to its overall high efficiency and cost-effectiveness⁵. Several microorganisms have diverse metabolic capabilities to degrade and utilise petroleum hydrocarbons as their carbon and energy sources via aerobic or anaerobic pathways^{6–8}.

The focus of the biodegradation of aliphatic and aromatic hydrocarbons has been on using a special community of microorganisms, which have the genes of interest⁹. To date, only scant information is available regarding the interaction of microbial functions and their metabolic diversity during biodegradation by various microbial populations in any affected environment¹⁰. The introduction of metagenomics can play a vital role in unearthing and monitoring the microbial communities by providing access to the taxonomic and functional gene composition¹⁰. According to Gong *et al.*¹¹, most of the metagenomic analysis tools have opened new windows of opportunities for researchers to analyse the microbial community as a whole (whole-genome sequencing) and the genetic diversity, which facilitates active metabolic pathways in any given environment.

While studies in the past have focused on species composition of a community, metagenomic studies (a gene-centric approach) enable assessment of the biological function of the gene rather than the taxonomic identity. The study of microbial biodiversity and whole-genome sequencing (WGS) and analysis of any complicated samples with numerous microorganisms, which are not pure and unculturable in any given laboratory, have made shotgun metagenomics a more efficient tool than most of the already existing conventional techniques⁴. Furthermore, there is always a well-established direct correlation between microbial diversity and the presence of hydrocarbons in any oil-contaminated sites. These studies have reported microbial species richness and diversity in different geographical locations. Furthermore, large variations have been observed in the functional capacity of the microbial communities identified at each location. Metagenomic studies therefore provide excellent opportunities for finding new microbial strains and genes involved in bioremediation of hydrocarbon contaminants. Therefore, the aim of this study was to assess the taxonomic profile of the metagenome and the microbial adaptations in two different petroleum-contaminated sites using a metagenomics approach.

Materials and Methods

Sample collection. The study was carried out between June and July 2016 and the petroleum-contaminated water samples were collected from two different petroleum-polluted water aquifers in Gezina (S -25.7°17'28.8, E 28.2°05'75.8), a suburb of Pretoria, South Africa. This is the only area in Pretoria where high-molecular-weight hydrocarbon oil spills or leakages from petroleum stations are discharged into water aquifers. Contaminated water samples, which were collected in 1 L sterile plastic bottles, were transported from the affected sites to the Tshwane University of Technology laboratory under standard refrigerated (−20 °C) conditions. To ensure the quality of collected water samples, the petroleum-contaminated water samples were thoroughly mixed prior to use.

DNA extraction. The samples were filtered using filter paper with a pore size of 0.45 µm (Whatman Filter Paper No. 1). Two hundred microlitres (200 µL) of PBS together with 750 µL of lysis solution were added to the residue and thoroughly mixed. The metagenomic DNA was extracted independently for all the samples with ZR fungal/bacterial DNA Kit (Zymo Research), as per manufacturer's instruction. Finally, the DNA was eluted in 30 µL of ultrapure water (Milli-Q water). The integration of the metagenomic DNA was assessed on the 0.8% agarose gel. The DNA quantity, quality and purity were measured using a NanoDrop ND-1000 spectrophotometer (PeqLab, Germany) where a 260/280 ratio of ~1.8 was accepted as “pure” for DNA.

PCR amplification. The PCR amplification was performed on the extracted DNA samples using a combination of the universal 18S rRNA gene forward primer 5'-GACGGGCGGTGTGTACA-3' and the reverse primers 5'-CTGGTTGATCCTGCCAG-3' or 5'-TGATCCTTCYGCAGGTTTAC-3', respectively¹². These primers were used to amplify approximately 500 bp of the 18S rRNA gene sequence. For the final reaction, each PCR reaction mixture (50 µL) contained DreamTaq™ Green DNA polymerase (dNTPs and 4 mM MgCl₂), Master Mix (2×) (25 µL), nuclease-free water (22 µL), forward primer (1 µL) (0.2 µM) and reverse primer (1 µL) (0.2 µM) and metagenomic DNA (50–100 ng). The PCR cycling parameters used were as follows: initial denaturation step at 95 °C for 5 min, followed by 30 cycles of 94 °C for 1 min, 55 °C for 30 s and 72 °C for 1 min 30 s, with a 10 min extension at 72 °C, that was finally followed by rapid cooling to 4 °C. The PCR products were loaded into a 1% agarose gel and then visualised with an ultraviolet (UV) transilluminator (InGenius Bio Imaging System, Syngene, Cambridge, UK). The DNA was finally re-amplified with same sets of the two primers. The pooled samples were sequenced on the GS-FLX-Titanium series 454/Roche by Inqaba Biotechnology, South Africa.

Sequencing of genomic DNA. Genomic DNA was sent to Inqaba Biotechnical Industries, a commercial NGS service provider, for sequencing. Briefly, genomic DNA samples were PCR amplified using a universal primer pair (566 F and 1200 R - targeting the 18S rRNA gene). Resulting amplicons were gel purified, end repaired and illumina specific adapter sequence were ligated to each amplicon. Following quantification, the samples were individually indexed, and another bead based purification step was performed. Amplicons were then sequenced on illumina's MiSeq platform, using a MiSeq v3 (600 cycle) kit. 20 Mb of data (2 × 300 bp long paired end reads) were produced for each sample. The BLAST-based data analysis was performed using an Inqaba in-house developed data analysis pipeline.

Statistical analysis. The data were statistically analysed using the Stata computer software (version: STATA V12, STATA Corp. LP, 2012). Analysis of variance (ANOVA) was performed to compare the average percentage abundance between different microbes in two petroleum spill sites.

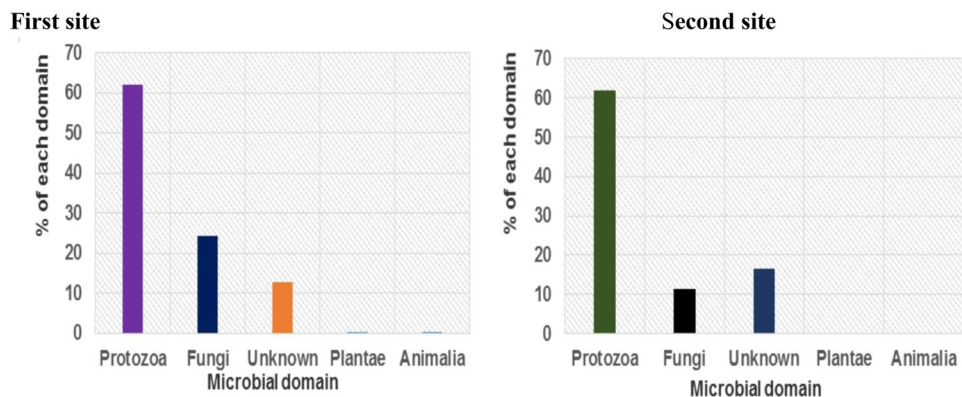


Figure 1. The different microbial domains observed in the first and second sites in petroleum-polluted water aquifers.

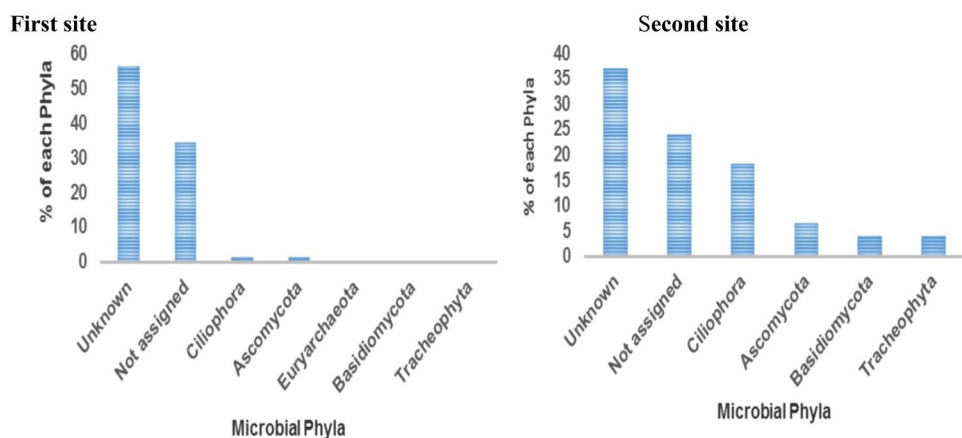


Figure 2. The different microbial phyla observed in the first and second site in petroleum-polluted water aquifers.

Results

Domain-level microbial communities in petroleum oil-contaminated water aquifers. The various domains of microbes from two different petroleum oil-contaminated water sites are clearly illustrated in Fig. 1. In the first site, protozoa (62.04%) were found to be the most dominant group, followed by fungi (24.49%), unknown (12.87%) and finally other sequences such as Animalia and plantae which were below (0.10%). In the second site, the microbial communities were identified in the following order: protozoa (61.90%), unknown (16.51%) and fungi (11.41%) domains (Fig. 1). Statistically, there was no significant difference at $p < 0.05$ in terms of the domains of microbes observed in these two sites.

Phylum-level microbial diversity of petroleum oil-polluted water sites. A representation of phylum-level microbial diversity identified in two petroleum-contaminated sites is depicted in Fig. 2. The dominant groups were found to be the unknown (56.45%) and the unclassified eukaryotic lineages derived from Eukaryota (34.55%), while the remaining phyla were considered insignificant (Fig. 2). These observations were similar to those of the second site where unknown sequences at the genus level derived from the phylum Eukaryota occupied the first position (37.14%), followed by unclassified sequences from phylum Eukaryota (24.07%), Ciliophora (18.17%) and Ascomycota (6.47%) (Fig. 2). A marked reduction in the quantitative number of the Tracheophyta (4.01%) and other phyla was observed in the second site (Fig. 2). Statistically, there was no significant difference at $p < 0.05$ in terms of the microbial phyla observed in these two sites.

Taxonomic classification of microbes in petroleum oil-polluted water sites at the genus level. According to the classification at the genus level, the dominant group was *Naegleria* (15.21%) followed by *Vorticella* (6.67%) as the only ciliated protozoan genus, other species such as *Arabidopsis* (2.97%), *Asarum* (1.84%) *Populus* (1.04%) were significantly low and drastically lower in the first site (Fig. 3). Regarding the second site, the dominant group was *Naegleria* (18.29%) followed by *Colpoda* (9.86%) with the remainder of the genera representing $< 2\%$ (Fig. 3). It was evident that there were few identified protozoan genus in this site due to the high number ($> 0\%$) of unclassified sequences derived from the phylum Eukaryota (Fig. 2).

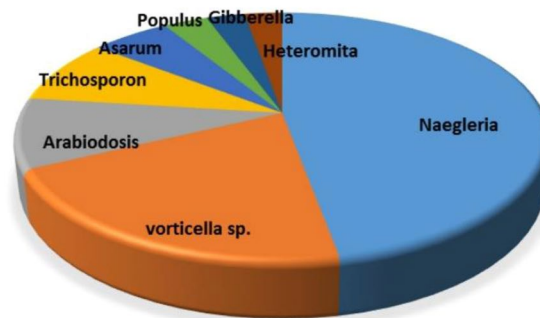
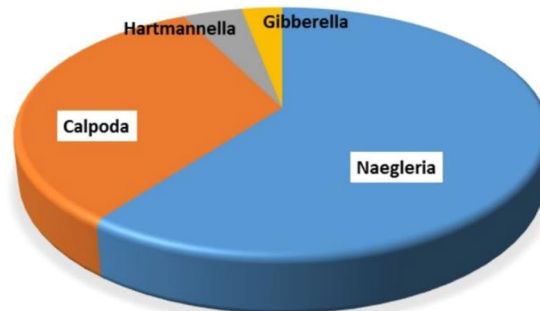
First site**Second site**

Figure 3. Different genera of the microbes observed in the first and second sites in petroleum-polluted water.

Discussion

Petroleum hydrocarbon contamination of the existing freshwater sources and the environment at large is of growing public concern as large amounts of these compounds are being released deliberately or through accidental spills¹². Petroleum oil exploration has led to a massive release of huge amounts of oil into soil and water and the majority of these hydrocarbons are usually toxic to the end-users, as reported by Jin *et al.*¹³. In spite of these negative impacts, there are also positive aspects where the affected sites contain a diverse array of microbes, which possess unique enzymes or genes for the biodegradation of hydrocarbon contaminants⁴. According to Abbasian *et al.*¹⁴, the application of metagenomic analysis tools in the affected oil-spill sites provides a promising approach that allows analysis of these microbial communities and their mode of adaptation to petroleum oil contamination in the environment.

Results of the present study have demonstrated that the microbial community of the petroleum oil-polluted water aquifers of the two different sites in Pretoria/South Africa consisted of various microbial domains with the protozoa (eukaryotes) were the most dominant groups, followed by fungi and archaea which were present in low numbers in both oil-polluted aquifer sites, which were present in low numbers in both oil-polluted aquifer sites. The results obtained in this study are in line with the findings of previous studies done by Liang *et al.*¹⁵ and Yergeau *et al.*¹⁶ on metagenomic analysis of the petroleum-contaminated water obtained from the northern parts of North America and southern parts of China. According to Andrews and Floodgate¹⁷, protozoan microbes such as ciliates and heterotrophic dinoflagellates are mainly consumers of phytoplankton and they are also important contributors to their diet; hence, they are also able to ingest oil droplets and oil-contaminated phytoplankton. Bioaccumulation of hydrocarbon components of petroleum oil do stimulate higher uptake by protozoans and this process will therefore facilitate and enhance removal of oil from polluted water, thereby reducing the availability of oil¹⁸. Nevertheless, the potential role of protozoans in the interactions between dispersed petroleum oil and other aquatic microbes (biomagnification or mitigation) has generally been neglected in petroleum toxicological and bioaccumulation studies in petroleum-contaminated sites as further reported by Calbet¹⁸.

The representation of different phyla of microbes obtained in two petroleum-contaminated aquifer sites showed the dominant phyla to be the Eukaryota [(where a high number (>20%) of unclassified sequences were observed)] (Fig. 2). According to Atlas¹⁹ and Widdel *et al.*²⁰, the majority of Eukaryotic family are reported to be more closely monitored in order to evaluate the potential for both aerobic and anaerobic oil biodegradation since they are known for their utilisation of petroleum hydrocarbons. Most eukaryotic organisms which are distant cousins of the protozoans such as the Ascomycota are mainly ciliates and flagellates. Kiørboe *et al.*²¹ pointed out that ciliates are common grazers in many aquatic environs and have been reported in many sites involving oil degradation or hydrocarbon-contaminated aqueous habitats. Parker²², Batelle²³ and Andrews and Floodgate¹⁷ have reported that the protozoan phyla are well adapted to graze upon oil-degrading communities and to survive in severely polluted habitats. They reproduce rapidly and show a wide tolerance against environmental extremes in temperature, oxygen supply, pH and the presence of toxins. The phylum Ascomycota plays an important role in removing hazardous hydrocarbon compounds from water

and petroleum oil from oil spills. This is the desired ecological niche for fungi, which inhabit substrates, and in biodegradation processes they are known as consumers of the carbon sources from hydrocarbons found in existing freshwater sources¹⁰. According to Ojo²⁴, fungi have been found to be better degraders of petroleum than traditional bioremediation techniques including bacteria, and although hydrocarbon degraders may be expected to be readily isolated from a petroleum oil-associated environment, the same degree of expectation may be anticipated for microorganisms isolated from a totally unrelated environment. This phylum is also found to be capable of biodegrading a structurally diverse range of oil-derived compounds by employing a variety of mechanisms including intracellular enzymes such as cytochrome P450 monooxygenases and transferases as well as extracellular enzymes such as laccases and fungal peroxidases as reported by Harms *et al.*²⁵. *Vorticella* (6.67%) was the only ciliated protozoan genus observed during this study period especially in the first site while none of the protozoa were found in the second site. This could be as a result of the high number (>30%) of unclassified sequences derived from the phylum Eukaryota in the second site which were not been able to be fully identified (Fig. 2).

Conclusion

Based on the metagenomic data analysis, the two petroleum-contaminated sites were inhabited by a diverse array of eukaryotic (protozoa) organisms were the predominant groups, followed by fungi and other domain such as plantae which were found to be present in relatively low numbers in both sites. These microorganisms play an important role in the bioremediation process as they are found to possess rich sources of different and diverse enzymes useful for the degradation of different hydrocarbon groups; in addition, most of these microorganisms secrete surfactants that facilitate emulsification of the petroleum oil, thus enhancing the uptake of hydrocarbons into the microbial cells. The novelty of this study is that the majority of these microbes were found to be protozoa (eukaryotes) found in two sites. It is envisaged that protozoans will form part of the bioremediation process. This study thus confirms the presence of protozoan phyla, which were never before known to be playing any role in hydrocarbon degradation or inhabiting these petroleum oil-contaminated areas. The phylogenetic novelty observed during this study makes an important contribution towards a better understanding of the microbial diversity that exists in an oil-spill affected site. The findings of this study provide a scientific background and technology platform for further research and utilisation of microbial species, especially protozoan species, in petroleum-contaminated areas in South African water sources and beyond. This study also allows capacity building for water-scarce countries in developing a high level of scientific research in the area of metagenomics.

References

1. Haritash, A. K. & Kaushik, C. P. Biodegradation aspects of polycyclic aromatic hydrocarbons (PAHs): a review. *J. Haz. Mater.* **169**(1–3), 1–15 (2009).
2. Van Hamme, J. D., Singh, A. & Ward, O. P. Recent advances in petroleum microbiology. *J. Micro. Mol. Bio. Rev.* **67**, 503 (2003).
3. Juhasz, A. L. & Naidu, R. Bioremediation of high molecular weight polycyclic aromatic hydrocarbons: a review of the microbial degradation of benzo[a]pyrene. *J. Inter. Biodeter. Biodegr.* **45**, 57–88 (2000).
4. Abbasian, F., Lockington, R., Mallavarapu, M. & Naidu, R. The integration of sequencing and bioinformatics in metagenomics. *Rev. Env. Sci. Biotechnol.* **14**, 357–383 (2015a).
5. Cerniglia, C. E. & Sutherland, G. R. Degradation of polycyclic aromatic hydrocarbons by fungi. In: Timmis, K. N. (Ed.), *Handbook of Hydrocarbon and Lipid Microbiology*, Springer, Berlin; London. 2079–2110 (2010).
6. Magot, M., Ollivier, B. & Patel, B. K. C. Microbiology of petroleum reservoirs. *Mikr.* **75**, 82 (2000).
7. Borzenkov, I. A., Milekhina, E. T., Gotoeva, M. T., Rozanova, E. P. & Beliaev, S. S. The properties of hydrocarbon-oxidizing bacteria isolated from the oilfields of Tatarstan, Western Siberia, and Vietnam. *Mikr.* **75**, 82–89 (2006).
8. Da Cruz, G. F. *et al.* Could petroleum biodegradation be a joint achievement of aerobic and anaerobic microorganisms in deep sea reservoirs? *AMB Exp.* **1**, 47–124 (2011).
9. Uhlík, O., Strejček, M., Hroudová, M., Demnerova, K. & Macek, T. Identification and characterization of bacteria with bioremediation potential: From cultivation to metagenomics. *Chem Listy.* **107**, 614–622 (2013).
10. Thomas, T., Gilbert, J. & Meyer, F. Metagenomics – a guide from sampling to data analysis. *Micro. Informat. Experi.* **2**, 3 (2012).
11. Gong, J. S. *et al.* Metagenomic technology and genome mining: emerging areas for exploring novel nitrilases. *Appl. Microbiol. Biotechnol.* **97**, 6603–6611 (2013).
12. Abbasian, F., Lockington, R., Mallavarapu, M. & Naidu, R. A. Comprehensive review of aliphatic hydrocarbon biodegradation by bacteria. *J. Appl. Biochem. Biotechnol.* **176**, 670–699 (2015b).
13. Jin, H. M., Kim, J. M., Lee, H. J., Madsen, E. L. & Jeon, C. O. *Alteromonas* as a key agent of polycyclic aromatic hydrocarbon biodegradation in crude oil-contaminated coastal sediment. *J. Env. Sci. Technol.* **46**, 7731–7740 (2012).
14. Abbasian, F. *et al.* Microbial diversity and hydrocarbon degrading gene capacity of a crude oil fields soil as determined by metagenomics analysis. *J. Biotechnol. Prog.*, <https://doi.org/10.1002/btpr.2249> (2016).
15. Liang, Y. *et al.* Functional gene diversity of soil microbial communities from five oil-contaminated fields in China. *J. Int. Society Microb. Ecol.* **5**, 403–413 (2011).
16. Yergeau, E., Sanschagrin, S., Beaumier, D. & Greer, C. W. Metagenomic analysis of the bioremediation of diesel-contaminated Canadian high arctic soils. *PLoS One* **7**, 1–10 (2012).
17. Andrews, A. R. & Floodgate, G. D. Some observations on the interactions of marine protozoa and crude oil residues. *Mar. Bio.* **25**, 7–12 (1974).
18. Calbet, A. The trophic roles of microzooplankton in marine systems. *Ices J. Mar. Sci.* **65**, 325–331 (2008).
19. Atlas, R. M. Microbial-degradation of petroleum-hydrocarbons – an environmental perspective. *Microb. Rev.* **45**, 180–209 (1981).
20. Widdel, F. & Grundmann, O. Biochemistry of the anaerobic degradation of non-methane alkanes. In: Timmis, Kenneth N., McGenity, Terry J., Vander Meer, Jan Roelof, De Lorenzo, Victor (Eds), *Handbook of Hydrocarbon and Lipid Microbiology*, Springer Verlag, Berlin, 909–924 (2010).
21. Kjørboe, T., Tang, K., Grossart, H. P. & Ploug, H. Dynamics of microbial communities on marine snow aggregates: colonization, growth, detachment, and grazing mortality of attached bacteria. *J. Appl. Env. Microb.* **69**, 3036–3047 (2003).
22. Parker, J. G. Ciliated protozoa in marine pollution studies. *Ecotox. Env. Safety.* **7**, 172–178 (1992).
23. Batelle, C. D. Mushrooms: Higher macro-fungi to clean up the environment. *Env. Issues Fall.* 2000 (2000).
24. Ojo, O. A. Petroleum-hydrocarbon utilization by native bacterial population from a wastewater canal Southwest Nigeria. *Afr. J. Biotechnol.* **5**, 333–337 (2005).
25. Harms, H., Schlosser, D. & Wick, L. Y. Untapped potential: exploiting fungi in bioremediation of hazardous chemicals. *Nat. Rev. Microb.* **9**, 177–192 (2011).

Author Contributions

Main authors Kachienga Leonard and K. Jitendra are the main contributor of this manuscript e.g. writing it up and performing the whole study. Second author Momba Maggy is only the scientific mentor of this manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018