RESEARCH ARTICLE

Metagenomic evaluation of a Utah tar sand microbiota suggests the predominant hydrocarbonoclastic role of Actinobacteria [version 1; peer review: 1 approved, 1 approved with reservations]

Dawn E. Lewis¹, Ashish Pathak², Cynthia B. Jones³, Charlemagne Akpovo⁴, Ashvini Chauhan⁵

¹Department of Biology, College of Science and Technology, Florida A&M University, Tallahassee, FL, 32307, USA
²School of the Environment, Environmental Biotechnology Laboratory, Florida A&M University, Tallahassee, FL, 32307, USA
³School of Pharmacy, Marshall University, Huntington, WV, 25755, USA
⁴Department of Physics, College of Science and Technology, Florida A&M University, Tallahassee, FL, 32312, USA
⁵School of the Environment, Environmental Biotechnology Laboratory, Florida A&M University, Tallahassee, FL, 32307, USA

First published: 17 Oct 2018, 7:1650
https://doi.org/10.12688/f1000research.16126.1
Latest published: 17 Oct 2018, 7:1650
https://doi.org/10.12688/f1000research.16126.1

Abstract
Background: Occurrences of tar sands have been reported in 22 states in the United States; however, the largest deposit is located in southwestern Utah. It has been suggested that tar sands were created by the microbial degradation of immobile subsurface oil over several million years; however, little is known about the indigenous microbial communities in the bituminous tar sands. Methods: This study identified Utah tar sand microbiota using next-generation sequencing technology and characterized the functional diversity using community-level physiological profile (CLPP). Results: Microbiota identified in these tar sands are mainly affiliated with the Gram-positive Actinobacteria and representatives of genera that have also been previously shown to degrade aromatic hydrocarbons, including Arthrobacter, Dietzia, Janibacter, Nocardioides, Microbacterium, Agrococcus and Salinibacterium, suggesting that these microbes likely play roles in the biodegradation of oil-hydrocarbons. CLPP analysis revealed less than 24 h was needed for the first color development in the microplate wells containing the polymers, whereas the duration of the lag phase of the carboxylic acids was prolonged. Conclusions: The quick utilization of the polymers suggests that the indigenous microbial community, especially the actinomycetes in the tar sand habitat, are poised and primed to degrade these recalcitrant compounds.
Introduction

‘Tar sands’ is a term used to describe a combination of bituminous sand, sandstone, oil-impregnated rock, oil sand and rock asphalt. The tar sands is a rapidly developing source of unconventional petroleum. The bitumen, an oil-rich residue, can be extracted from the tar sands and refined into crude oil. In the United States, occurrences of tar sands have been reported in 22 states; however, one of the largest deposit is located in southwester Utah, with estimated recoverable oil ranging from 12 to 20 barrels.

It has been suggested that tar sands were created by the microbial degradation of immobile subsurface oil over several million years. Little is known about the origin of the microorganisms in the tar sands of Utah, although the indigenous microbial communities exist in extreme conditions in the bituminous tar sands and are limited by harsh conditions such as low moisture and oxygen, recalcitrant hydrocarbons and a high concentrations of toxic metals. However, a paucity of data exists on the indigenous microbial communities of the tar sands of Utah.

Traditionally, the diversity of the bacterial communities in the tar sands has been investigated using the isolation and cultivation approaches. However, only 1.5% of bacteria in the soil can be readily cultured. Most cultural conditions cannot mimic the specific microhabitats that many prokaryotes thrive within. Therefore, new methods for the analysis of whole microbial community structure and metabolic function in the bituminous tar sands have been developed. One of the methods developed to provide a dynamic tool for the assessment of microbial community structure and function is the BIOLOG™ Community Level Physiological Profile (CLPP). This method is based upon the preferential metabolism of 31 different carbon sources on a microtitre plate. These carbon sources include a wide range of chemical classes, such as carbohydrates, esters, polymers, carboxylic acids, alcohols, amines, phosphorylated chemicals, amino acids, aromatic chemicals, and amines. Each well on the BIOLOG™ EcoPlate contains a unique carbon source, peptone, and 2,3,5-triphenyltetrazolium chloride (TTC) dye. NADH is produced from the respiration of the specific carbon sources. The final electron acceptor TTC is irreversibly reduced to formazan, a red pigment that can be quantified visually by use of a microplate reader. The intensity of the color change correlates to the amount of metabolism of the carbon source in that well. The net intensity of the color change is calculated by subtracting the absorbance of the non-carbon-source control well. The oxidation of the carbon substrate oxidized by the microbe can be considered to be its metabolic fingerprint. However, Windig reported that the BIOLOG™ system excludes strict anaerobes and bacteria that lack certain electron transport enzymes. Yao et al. advocated that CLPP is selective and favors microbes that grow quickly or those with a high inoculum density in the initial sample. Hence, culture-independent analysis is the method of choice for the investigation of the bituminous tar sands.

The objective of this study was to identify and characterize the indigenous bacterial communities of the Utah tar sands using next-generation sequencing technology and to assess the functional diversity using the CLPP system. The overall goal of this project is significant because studies have demonstrated that naturally occurring microbes can be harnessed for the degradation of recalcitrant polyaromatic hydrocarbons, heavy metals, and naphthenic acids. The combination of the functional diversity and the characterization of the indigenous microbiota will advance our understanding of the fate of tar-associated potentially toxic compounds by environmental microbiota.

Methods

Microbial community analysis by genomic DNA extraction and 454 pyrosequencing

The University of Alberta Geotechnical Centre, Canada, kindly provided the tar sands sample. All samples were shipped on ice to Florida A&M University where they were processed for metagenomics and CLPP analyses. Genomic DNA was extracted from the samples using the PowerSoil DNA Isolation kit (MoBio Laboratories, Inc., Carlsbad, CA, USA). Bar-coding pyrosequencing of the 16S rRNA gene V4 region was performed using the Earth Microbiome Project (EMP) standard protocols (http://www.earthmicrobiome.org/emp-standard-protocols) and sequenced on a Roche 454 FLX instrument (Roche, Indianapolis, IN, USA) with titanium reagents, following manufacturers recommended procedures. Sequences that passed quality controls were uploaded to MOTHUR v.1.33.3, where tags, low-quality sequences and chimeric reads were removed.

CLPP analysis

CLPP of the tar sand sample was analyzed using BIOLOG™ Ecoplates, as described previously. Triplicate samples of the BIOLOG™ Ecoplate were used for this experiment. A total of three sets of 1 g samples of the tar sand were shaken in 99 ml of distilled sterile water for 20 min at 20°C. The samples were then incubated at 4°C for 30 min and then centrifuged for 10 min at 500g. The supernatant of the tunes were pooled and thoroughly mixed. In total, 150 μl of the supernatant was inoculated into each well of the Biolog EcoPlates. This ensured that each of the 96-well plate contained three replicates of each sample exposed to 31 different carbon sources. Water blanks were included with each plate and the plates were incubated at 25°C without shaking. Substrate utilization was monitored by measuring absorbance at 595 nm using a BIOTEK UQuant Microplate Spectrophotometer, (BIO-TEK Instruments, Inc., USA). The first measurement was made immediately after inoculation and subsequent readings were taken after every 24 h interval for 7 days. The measurements of individual substrates were corrected for background absorbance by subtracting the absorbance of the control (water) samples. If a negative number was obtained, it was manually set to zero. A well was considered positive if the mean OD exceeded that of 0.25.

In this study the 31 carbon sources were organized into groups 1–5 as described by Zak et al. Carbon sources originally grouped as miscellaneous by Zak et al. were those new to the BIOLOG™ EcoPlate, were grouped into one of the other five categories - carbohydrates, carboxylic acids, amino acids, esters and
polymers. Grouping the data into 5 guilds compresses a 31-dimensional data set into 5 dimensions, significantly reducing the complexity of the data and subsequent interpretation\(^a\).

The average well color development (AWCD) was used to assess the microbial response for all the carbon sources. The AWCD was determined as follows:

\[
\text{AWCD} = \frac{\sum \text{OD}_{i}}{31}
\]

where \(\text{OD}_i\) is optical density value from each well, corrected subtracting the blank well (inoculated with water) values from each plate well\(^b\). The 96-h data was used for statistical analysis of CLPPs.

The diversity of substrate utilization was calculated using the Shannon’s diversity index \((H)\) and evenness \((E)\)^\(^c\). The Shannon-Weaver index is a measure of the capacity of the bacterial community to degrade the carbon sources in the well and can be considered as an index of the physiological diversity of the bacterial community in the tar sands. Higher values of \(H\) indicate the ability of the microbial community to degrade the substrates with a high efficiency\(^d\):

\[
H = -\sum p_i \ln p_i
\]

where \(p_i = (\text{OD reading of well } i)/(\text{sum of all wells})\), based on the OD in the Ecoplates.

The Shannon evenness \((E)\) is a measure of how dissimilar the abundances of the species in a community are from each other\(^e\) and was calculated as:

\[
E = H / \ln S
\]

where \(S\) (Shannon richness) is the number of substrates used by the microbial communities. Further, the cluster analysis of the substrate utilization pattern was constructed using the nearest neighbor method with Euclidian distance to form linkage dendrogram.

**Results and discussion**

**Microbial community analysis of tar sands**

In this study, the compositions of bacterial community of the tar sands was investigated by a pyrosequencing-based analysis of the 16S rRNA gene sequences. The gram-positive Actinobacteria (50%) were dominant in the tar sand followed by Betaproteobacteria (27%), Alphaproteobacteria (7%), Gammaproteobacteria (7%) and Acidobacteria (2%) (Figure 1).

![Figure 1. Taxonomic distribution of the obtained metagenome sequences shown at the phylum level. The gram-positive Actinobacteria were dominant in the tar sand followed by Betaproteobacteria, Alphaproteobacteria, Gammaproteobacteria and Acidobacteria.](image-url)
Actinobacteria are known for their role in the biodegradation of a variety of different pollutants including petroleum hydrocarbons\textsuperscript{18,19}. The predominant genus identified in these tar sands was \textit{Arthrobacter}, followed by \textit{Dietzia}, \textit{Janibacter}, \textit{Nocardioides}, \textit{Microbacterium}, \textit{Agrococcus} and \textit{Salinibacterium} (Figure 2). Most of these genera have also been previously shown to degrade aromatic hydrocarbons, indicating that tar sands are a very active repository of hydrocarbonoclastic microorganisms. When the taxonomic affiliation of the obtained metagenomic sequences were investigated, we found that the gram-positive \textit{Arthrobacter} spp. from the Actinobacteria phyla were predominant, which comprised approximately half of the total microbial community assemblage in this particular Utah tar sand. \textit{Arthrobacter} species possess a significant hydrocarbonoclastic potential as demonstrated by previous studies and we recommend that Actinobacteria, native to the tar sand habitats, should be targeted for future research on this area.

**Microbial community metabolic analyses**

Biolog\textregistered Ecoplates were used to evaluate the functional diversity of the microbial community of the tar sands. The AWCD revealed sigmoidal relationships between the OD\textsubscript{590} and time for all carbon sources (polymers, carbohydrates, amino acids, esters) except the carboxylic acids (Figure 3). This pattern
for the polymers, carbohydrates, amino acids, and esters is fundamentally similar to the bacterial growth curve. Yao et al. suggested that color development reflected species metabolic activity and the ability of the bacterial community to respond to substrates. The AWCD for all the substrates in the Biolog® Eco-microplates was 0.25. Less than 24h was needed for the first color development the microplate wells containing the polymers whereas the duration of the lag phase of the carboxylic acids was prolonged. The first color development of the carboxylic acids (AWCD = 0.2) was recorded at 48 h (Figure 3). The highest utilization was observed in the microplate wells containing the polymers (AWCD = 0.78) after 96 hours. The observed differences in the substrate utilization pattern of five major carbon groups (Figure 3) might be due to the presence of different functional groups, such as carbohydrates (R-C=O), amino acids (–NH₂ and –COOH), carboxylic acid (–COOH), esters, amines and phosphorylated acids (–COOR’, –NH₂), and polymers (–(CH₂–CH₂)n(R)). The high utilization of the recalcitrant polymers suggest that the polymers could be more easily utilized by the indigenous microbial community in the tar sands. After 96 h, there was a decrease in the rate of AWCD; thus, the analysis of functional diversity of microbial community of tar sands was completed at 96 h. The decreased rate may be due to the change of metabolic activity of active microbial communities utilizing the C-substrates. Cluster analysis of the substrates revealed a systematic grouping of the substrates based on their utilization pattern (Figure 4).

The Shannon-Wiener index is an indication of the spread or distribution of carbon source utilization by the microbial community. Typical values of the Shannon-Wiener Index usually lies between 1.5 and 3.5, and rarely exceed 4.0. In this study, the Shannon-Weiner index, H, of the microbial community in the tar sands was 3.072 and the Shannon evenness, E, was 1.009. These values indicate a more diverse microbial community and an even distribution of the species within the tar sands.

### Dataset 1. Raw data obtained from the community-level physiological profile

[https://doi.org/10.5256/f1000research.16126.d218909](https://doi.org/10.5256/f1000research.16126.d218909)

### Conclusion

This study investigated the indigenous microbial communities of the tar sands of Utah in an effort to understand the structure and the functional diversity of the microbial community. The substrate utilization patterns of resident...
microbial communities and the identification of the hydrocarbonoclastic microorganisms in the tar sands provide valuable baseline information than can be used for hydrocarbon bioremediation and for devising biotechnological approaches to tar sands bioremediation efforts.

Data availability
Dataset 1. Raw data obtained from the community-level physiological profile, 10.5256/f1000research.16126.d21890923.

The DNA sequences from this metagenomic project are available from the Sequence Read Archive-European Nucleotide Archive, accession number SRR1699470.

References

1. Spieker EM: Contributions to economic geology (short papers and preliminary reports). 1930; Part II, Mineral fuel. Bituminous sandstone near Vernal, Utah. U.S. Geological Survey Bulletin. 1930; 822: 77–98. Publisher Full Text
2. Baxter SR: Tar Sands: Worth the Energy? An Analysis of the Future of Utah’s Tar Sands. J Land. Resources & Envr L. 2007; 27(2): 323–344. Reference Source
3. Hamer NK, Richardson TL, Thompson KA, et al.: Microbial processes in the Athabasca Oil Sands and their potential applications in microbial enhanced oil recovery. J Ind Microbiol Biotechnol. 2011; 38(11): 1761–75. Publisher Abstract | Publisher Full Text
4. Kim JS, Crowley DE: Microbial diversity in natural asphalts of the Rancho La Brea Tar Pits. Appl Environ Microbiol. 2007; 73(14): 4579–91. Publisher Abstract | Publisher Full Text | Free Full Text
5. Demeter MA, Lemire J, George J, et al.:Harnessing oil sands microbial communities for use in situ naphthenic acid bioremediation. Chemosphere. 2014; 97: 78–85. Publisher Abstract | Publisher Full Text
6. Amann RI, Ludwig W, Schleifer KH: Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol Rev. 1995; 59(1): 143–169. Publisher Abstract | Free Full Text
7. Head IM, Saunders JR, Pickup RW: Microbial Evolution, Diversity, and Ecology: A Decade of Ribosomal RNA Analysis of Uncultivated Microorganisms. Microb Ecol. 1998; 36(1): 1-21. Publisher Abstract | Publisher Full Text
8. Garland JL, Mills AL: Classification and characterization of heterotrophic microbial communities on the basis of patterns of community-level sole-carbon-source utilization. Appl Environ Microbiol. 1991; 57(6): 2351–2359. Publisher Abstract | Publisher Full Text
9. Bochner BR: Sleuthing out bacterial identities. Nature. 1989; 339(6220): 157–158. Publisher Abstract | Publisher Full Text
10. Winding A: Fingerprinting bacterial soil communities using Biolog microtitre plates. In: Ritz K, Dighton J, Giller K (eds.) Beyond the Biomass, Wiley and Sons, Chichester, 1994; 85–94. Reference Source
11. Yao H, He Z, Wilson MJ, et al.: Microbial Biomass and Community Structure in a Sequence of Soils with Increasing Fertility and Changing Land Use. Microb Ecol. 2000; 40(3): 223–237. Publisher Abstract
12. Insam H: A new set of substrates proposed for community characterization in environmental samples. In: Insam, H., Rangger, A. (Eds.), Microbial Communities. Springer-Verlag, Berlin/Heidelberg, Germany, 1997; 259–260. Publisher Full Text
13. Zak JC, Willig MR, Moorhead DL, et al.: Functional diversity of microbial communities: A quantitative approach. Soil Biol Biochem. 1994; 26(9): 1101–1108. Publisher Full Text
14. Weber KP, Legge RL: One-dimensional metric for tracking bacterial community divergence using sole carbon source utilization patterns. J Microbiol Methods. 2009; 79(1): 55–61. PubMed Abstract | Publisher Full Text
15. Weber KP, Grove JA, Gehder M, et al.: Data transformations in the analysis of community-level substrate utilization data from microplates. J Microbiol Methods. 2007; 63(3): 461–9. PubMed Abstract | Publisher Full Text | Free Full Text
16. Han XM, Wang RQ, Liu J, et al.: Effects of vegetation type on soil microbial community structure and catabolic diversity assessed by polyphasic methods in North China. J Environ Sci (China). 2007; 19(10): 1228–1234. PubMed Abstract | Publisher Full Text
17. Smith B, Wilson J: A consumer’s guide to evenness indices. Oikos. 1996; 76(1): 70–82. Publisher Full Text
18. Yeung CW, Woo M, Lee K, et al.: Characterization of the bacterial community structure of Sydney Tar Ponds sediment. Can J Microbiol. 2011; 57(6): 493–503. Publisher Abstract | Publisher Full Text
19. Pathak A, Green SJ, Ogram A, et al.: Draft Genome Sequence of Rhodococcus opacus Strain M213 Shows a Diverse Catabolic Potential. Genome Announc. 2013; 1(1): pii: e00144-12. PubMed Abstract | Publisher Full Text | Free Full Text
20. Yao HY, Xu JM, Huang CY: Substrate utilization pattern, biomass and activity of microbial communities in a sequence of heavy metal-polluted paddy soils. Geoderma. 2002; 115(1–2): 139–148. Publisher Full Text
21. Harch BD, Correll RL, Meech W, et al.: Using the Gini coefficient with BIOLOG substrate utilisation data to provide an alternative quantitative measure for comparing bacterial soil communities. J Microbiol Methods. 1997; 28(1): 81–101. Publisher Full Text
22. May RM: Patterns of species abundance and diversity. In Ecology and Evolution of Communities (ed. M. L. D. Cody, J. M.). Cambridge, Mass.: Harvard University Press. 1975. Reference Source
23. Lewis DE, Pathak A, Jones CB, et al.: Dataset 1 in: Metagenomic evaluation of a Utah tar sand microbiota suggests the predominant hydrocarbonoclastic role of Actinobacteria. F1000Research. 2018. http://www.doi.org/10.5256/f1000research.16126.d218909
Open Peer Review

Current Peer Review Status: ✔️ ?

Version 1

Reviewer Report 25 March 2020

https://doi.org/10.5256/f1000research.17609.r60547

© 2020 Vainstein M. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Marilene H. Vainstein
Center of Biotechnology, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

The manuscript "Metagenomic evaluation of a Utah tar sand microbiota suggests the predominant hydrocarbonoclastic role of Actinobacteria" describes the Utah sand microbiota using NGS and CLPP. Although the experiments are well-designed, the authors should boost the "Results and Discussion" section to create a more informative manuscript, exploring better the obtained results. Furthermore, some minor corrections are needed:

Figure 1: There are four Acidobacteria groups (Gp16, Gp7, Gp6, and Gp4). What is the difference between these groups?

Figure 1: Figure and data explanation through the text does not match the figure data. For example, the authors described that "The gram-positive Actinobacteria (50%) were dominant in the tar sand followed by Betaproteobacteria (27%), Alphaproteobacteria (7%), Gammaproteobacteria (7%) and Acidobacteria (2%)", but in the figure, Actinobacteria correspond to 48% and Betaproteobacteria 26%, while Alphaproteobacteria 8%.

"Actinobacteria are known for their role in the biodegradation of a variety of different pollutants including petroleum hydrocarbons\(^{18,19}\)." The authors can explore better this statement, maybe they can describe scientific and commercial applications of Arthrobacter, Dietzia, Janibacter, Nocardioides, Microbacterium, Agrococcus and Salinibacterium.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes
If applicable, is the statistical analysis and its interpretation appropriate?  
Yes

Are all the source data underlying the results available to ensure full reproducibility?  
Yes

Are the conclusions drawn adequately supported by the results?  
Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Biofilms, oil degradation, bioremediation.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 18 February 2019

https://doi.org/10.5256/f1000research.17609.r41171

© 2019 Huang C. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chunkai Huang  
Department of Civil and Environmental Engineering, University of Alberta, Edmonton, AB, Canada

This manuscript documented microbial communities in Utah tar sand. The methods are well described and results are clearly presented. The performance of these communities to consume 31 different carbon sources was compared. Conceptually, I like the idea to compare the degradation rate of 31 different carbon sources using Utah tar sand indigenous microbial communities at micro plate. However, I believe the microbial community characterization, especially the high-throughput data, are not vigorously examined. I have a few minor concerns as follows.

Major concerns:

1. As you mentioned in the Introduction part, the BIOLOG system excludes strict anaerobes and bacteria that lack certain electron transport enzymes. That means the indigenous microbes of Utah sand were selected after inoculation and cultivation at 96-well plates. Some of the Utah sand anaerobes and other microbes were inhibited. It is possible that those inhibited Utah sand anaerobes are efficient for the degradation of other recalcitrant compounds.

2. It is necessary to mention the primer pairs and the regions of the 16S rRNA genes, because it is essential to ensure full reproducibility by others.
3. All raw pyrosequencing data that was obtained from this study should be deposited to the NCBI Sequence Read Archive, and the accession no. should be provide in the method.

**Is the work clearly and accurately presented and does it cite the current literature?**
Yes

**Is the study design appropriate and is the work technically sound?**
Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**
Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**
Yes

**Are all the source data underlying the results available to ensure full reproducibility?**
Yes

**Are the conclusions drawn adequately supported by the results?**
Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Biological wastewater treatment, biofilms, oil sands process-affected water (OSPW)

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com