



ELSEVIER

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/mex

Method Article

Advanced methods for RNA recovery from petroleum impacted soils



Maria Irianni-Renno, Tom C. Sale, Susan K. De Long*

Department of Civil and Environmental Engineering, Colorado State University, 1301 Campus Delivery, Fort Collins, CO 80523, USA

A B S T R A C T

Microbially-mediated hydrocarbon degradation is well documented. However, how these microbial processes occur in complex subsurface petroleum impacted systems remains unclear, and this knowledge is needed to guide technologies to enhance microbial degradation effectively. Analysis of RNA derived from soils impacted by petroleum liquids would allow for analysis of active microbial communities, and a deeper understanding of the dynamic biochemistry occurring during site remediation. However, RNA analysis in soils impacted with petroleum liquids is challenging due to: (A) RNA being inherently unstable, and (B) petroleum impacted soils containing problematic levels of polymerase chain reaction (PCR) inhibitors that must be removed to yield high-purity RNA for downstream analysis. A previously published soil wash pretreatment step and a commercially available DNA extraction kit protocol were combined and modified to be able to purify RNA from soils containing petroleum liquids.

- A key modification involved reformulation of the pretreatment solution via replacing water as the diluent with a commercially-available RNA preservation solution.
- Methods were developed and demonstrated using cryogenically preserved soils from three former petroleum refineries. Results showed the new soil washing approach had no adverse effects on RNA recovery but did improve RNA quality, by PCR inhibitor removal, which in turn allows for characterization of active microbial communities present in petroleum impacted soils.
- In summary, our method for extracting RNA from petroleum-impacted soils provides a promising new tool for resolving metabolic processes at sites as they progress toward restoration via natural and/or engineered remediation.

© 2021 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

* Corresponding author.

E-mail address: Susan.De_Long@colostate.edu (S.K. De Long).

ARTICLE INFO

Method name: RNA purification from soils containing petroleum liquids

Keywords: RNA extraction, RNA purification, Petroleum liquids, Natural source zone depletion, Sequencing, RT-qPCR, Biodegradation

Article history: Received 24 March 2021; Accepted 28 August 2021; Available online 31 August 2021

Specifications table

Subject Area:	Engineering
More specific subject area:	Environmental engineering, microbiology, biogeochemical characterization of petroleum liquid impacted soils, remediation
Method name:	RNA purification from soils containing petroleum liquids
Name and Reference of Original Methods:	Pretreatment steps for the DNA extraction of two Dutch oilfield core samples. [1] van der Kraan, G.M., De Ridder, M., Lomans, B.P. and Muyzer, G. (2010). Sampling and nucleic extraction procedures from oil reservoir samples. <i>Applied microbiology and molecular biology in oilfield systems</i> , pp. 7–16, Springer. Dordrecht, 2010.
Resource availability:	Mo bio, Powerlyzer™ DNA kit (Mo bio, Carlsbad, CA.) n/a

Method details

Introduction

Broad realization has come to the fact that soils, comingled at a pore-scale with nonwetting petroleum liquids, are populated by microbes that can transform petroleum into CO₂. Surprisingly, rates of depletion of petroleum liquids via microbially-mediated processes commonly rival and/or exceed depletion rates achieved with engineering remedies [2,3]. These processes of converting petroleum liquids in soil-groundwater systems are referred to as natural source zone depletion (NSZD) per ITRC [4].

NSZD has been well-demonstrated through various lines of evidence, including quantification of CO₂ efflux through the vadose zone using a gas gradient method [5], the dynamic chamber method [6], and CO₂ traps [7]. Also, and perhaps more definitively, NSZD has been demonstrated by quantification of the heat generated from biotic oxidation of petroleum hydrocarbons in soil groundwater [8,9]. Biotic oxidation of petroleum hydrocarbons, like oxidation of organic composts, produces CO₂ and heat.

While the fact that NSZD happens is clear, the microbially-mediated biochemistry of NSZD, and how it might be enhanced, is not clear. A promising approach to better resolving the biochemistry of NSZD is characterization of microbial communities in soils impacted with petroleum liquids [10–12]. However, most hydrocarbon source zone research has characterized systems at the DNA level [2,12], only providing information on microbe presence. Since presence or absence of a genetic sequence does not reflect active (biodegrading) or even viable microbes, DNA-based analysis might paint an incomplete picture of processes driving hydrocarbon biodegradation at the time of sampling. In contrast, assessing a microbial community at the RNA level provides more direct information regarding active organisms at impacted sites. Herein, methods are advanced for RNA purification, quantification of gene expression, and sequencing in support of resolving the active (RNA-based) vs. the historical (DNA-based), metabolic processes driving NSZD. The challenge for RNA-based characterization of microbial ecology is the molecular instability of RNA, and petroleum-impacted sites present a unique challenge in this regard. Given instability, preservation during purification of RNA is critical.

With respect to RNA preservation at contaminant-impacted field sites, cryogenic collection of soil samples is promising [13,14]. The primary challenge with resolving active microbial ecologies in soils impacted by petroleum liquids is extraction and purification of high-quality RNA. Herein, a novel sample pretreatment procedure is advanced (“Step 2”, Methods section) for removing polymerase chain reaction (PCR) inhibitors from soils containing petroleum liquids, while preserving the integrity

of the RNA. Techniques are demonstrated using soil cores collected from three former petroleum refineries (Sites). The benefits of the new pretreatment wash method are demonstrated with samples from Site 1, spiked with *E. coli* as a source of RNA. Improved RNA quality and lack of inhibitors is further demonstrated with samples from Site 2. Lastly, samples from Site 3 are used to illuminate that subsurface microbial communities are dynamic, as evidenced by differences between the microbial phylotypes that were active at the time of sampling (RNA) and the microbial phylotypes that were present (DNA). Microbial data combined with other Site 3 characterization (unpublished data) provided evidence for a transition from historical anaerobically-mediated NSZD to an emerging, more aerobic, NSZD at the leading edge of a subsurface petroleum pool. Overall, documented RNA methods combined with DNA data provide a promising new tool for resolving the dynamic biochemistry of sites as they progress toward restoration via natural and engineered remedies.

Methods

The following describes our RNA extraction methods and demonstrative applications of our methods using soil core from three former refineries. In order to continue with downstream PCR based analysis, recovered RNA is transcribed to stable cDNA (a DNA copy of the recovered RNA), and cDNA is used to resolve RNA-based microbial communities.

RNA extraction procedures

Fig. 1 provides an overview of our extraction workflow and subsequent analyses. The following outlines key steps in RNA isolation from soil matrices impacted by petroleum hydrocarbons. The novel aspect of our method (Step 2) involves a modification of a previously published sample pretreatment wash [1] used to purify DNA from petroleum liquid impacted soils. Our methods include washing samples with sodium chloride, disodium-EDTA, TritonX-100, dehydrated skimmed milk and polydeoxinocinic-deoxycytidilic-acid (pdldc) to remove PCR inhibitors. Disodium-EDTA chelates metals, such as iron, so metals do not precipitate out of solution [15]. Iron is found at sites impacted by petroleum liquids and interferes with PCR-based reactions. Low concentrations of a mild detergent, such as TritonX-100, will not lyse cells but promotes petroleum desorption, allowing petroleum liquids to float at the top of the wash solution [16]. Dehydrated skimmed milk, in high salinity environments, binds biological macromolecules such as lipids and sugars [17]. Organic molecules and chelated metals in solution are removed from the soil mix via discarding the supernatant in each washing step, after centrifugation. Polydeoxinocinic-deoxycytidilic-acid (pdldc) is a nucleic acid surrogate and may interact with RNAses. Our modification consists of replacing deionized (DI) water as the wash solution diluent with LifeGuard™ solution (Qiagen, Germantown, MD), to protect RNA from degradation during inhibitor removal wash steps. In addition to preserving nucleic acid integrity, by inactivating RNAses, LifeGuard™ solution prevents new microbial growth (Qiagen, Germantown, MD). Referenced solutions C1-C6 are proprietary components of the PowerSoil Powerlyzer™ DNA isolation kit (Qiagen, Germantown, MD).

- (1) *Sample Collection and Preservation* – Attributes of soil sample collection sites and soil sample collection methods are presented in Table 1. All samples were shipped for analysis to Colorado State University (CSU).
- (2) *Sample Pre-treatment* – (A) Five-gram frozen soil samples were removed from the freezer and placed in 50 ml centrifuge tubes. (B) 80 ng of dehydrated skimmed milk (VWR, Radnor PA.), 10 µg of pdldc (Sigma-Aldrich, St. Louis, MO), and 1 ml of LifeGuard™ Soil Preservation Solution (Qiagen, Germantown, MD) were added to each sample. Pdldc and milk casein, present in the added dehydrated skimmed milk, act as adsorption competitors for nucleic acids. (C) The soil mixtures were vortexed with a Gennie-II vortex (Qiagen, German,MD) for one minute and were incubated on ice for an additional minute. (D) Next, pretreated soil samples were washed three times with three different solutions. (E) Between washes the samples were vortexed for 2 min, centrifuged at 13,000 rpm for 3 min in a Sorvall Legend XTR™ centrifuge (Thermoscientific, Ashville, NC), and the supernatant discarded. (F) During the first wash step, the reagents added

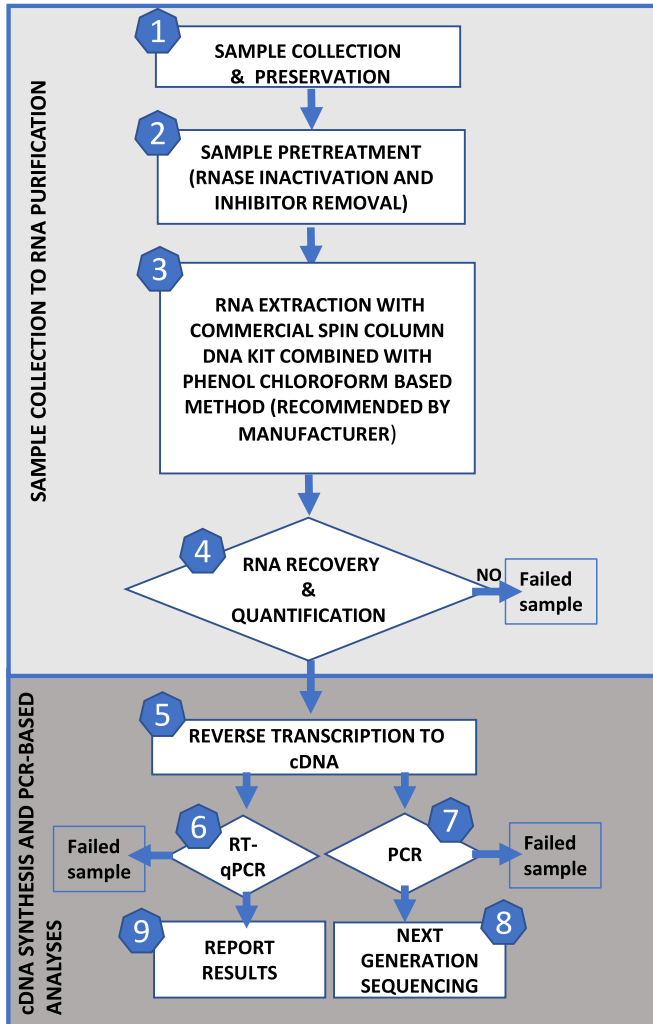


Fig. 1. Flow chart summarizing sample processes from sample collection to RNA purification, followed by cDNA synthesis and PCR based analyses.

to the mixture were: 8 ml of Lifeguard™ preservation solution, 500 µl of 50 mM tris-HCl (pH=8.3)(Sigma-Aldrich), 400 µl of 200 mM NaCl (VWR), 100 µl of 5 mM Na₂EDTA (Sigma-Aldrich), and 5 µl of Triton X-100 (5% V/V) (Sigma-Aldrich). (G) During the second wash step the following was added to the mixture: 9 ml of Lifeguard™ Soil Preservation Solution followed by the addition of 500 µl of 50 mM tris-HCl (pH=8.3), 400 µl of 200 mM NaCl, and 100 µl of 5 mM Na₂ EDTA. (H) The third washing solution contained 9.4 ml of Lifeguard™ Soil Preservation Solution, 500 µl of 50 mM tris-HCl (pH=8.3), and 100 µl of 5 mM Na₂ EDTA. After centrifugation and supernatant discard, the remaining soil solutions were ready for RNA extraction.

- (3) *RNA Extraction* – RNA was extracted using the PowerSoil Powerlyzer™ DNA isolation kit (Qiagen, Germantown, MD), via an alternative protocol with modifications adapted to isolate and purify RNA. In detail, (A) approximately 0.5 g of pretreated sample (or untreated for controls) were added to a dry bead tube from the commercial kit (compared to the

Table 1

Site attributes and sampling methods.

	Site 1	Site 2	Site 3
Site History	Crude oil refinery 1923 to 1982	Crude oil refinery 1904 to 1982	Crude oil refinery 1931 to 1986
Depositional Environment / Sediment	Braided stream channel deposit/Poorly sorted sands and gravel	Overbank flood plain deposit/ fine grained sand and silt	Glacial valley train deposit/ Poorly sorted sands and gravel
Primary Petroleum Liquids	Weathered gasoline and diesel	Weathered gasoline, diesel, and jet fuels	Weather gasoline, jet fuel, and diesel
Sampling Method	Direct push sampling of a single hole	Cryogenic coring (Kiaalhosseini et al., 2016)	Cryogenic coring per (Kiaalhosseini et al., 2016)
Samples and sample depth(s) below ground surface (bgs)	S1A – 2.4 m S1B – 0.2 m	S2 – 2.6 m	S3A, S3B, S3C – Samples collected from co-located triplicate boring from 9.3 m
Sample Preservation	Non-cryogenic direct push core collection. Sample immediately placed in a cooler with ice and shipped overnight to CSU. Stored at -20 °C until analyzed.	Liquid nitrogen-based cryogenic hollow stem auger core collection. Sample immediately placed in a cooler on dry ice in the field and shipped overnight to CSU. Stored at -80C° until analyzed.	
Total Petroleum Hydrocarbon Carbons in Soil (TPH)	S1A – 8500 mg/kg, 90% diesel range, 10% gasoline range S1B – <10 mg/kg	S2– 11,270 mg/kg 70% diesel range, 30% gasoline range	S3A – 12,917 mg/kg, 59% diesel range, 41% gasoline range S3B– 11,776 mg/kg, 51% diesel range, 49% gasoline range S3C – 8658mg/kg, 40% diesel range, 60% gasoline range

manufacturer recommended 0.25 g.) (B) Next the following were added to each sample: 500 μ l of the kit Bead Solution including 200 μ l phenol:chloroform:isoamyl alcohol (pH 7-8) (Amresco, Solon, OH), and 60 μ l of Solution C1. The addition of 200 μ l phenol: chloroform:isoamyl alcohol (pH 7-8) was a modification recommendation by the manufacturer (MoBio) and is not included as part of the protocol provided with the commercial kit. (C) The tubes were then vigorously shaken in a Powerlyzer™ (Qiagen, Germantown, MD.), according to the kit's provided protocol. (D) After 1 min of full-speed (13,000 rpm) centrifugation, solution C2 (200 μ l) and solution C3 (100 μ l) were added to the supernatant. (E) Each mixture was incubated for 5 min at 4 °C and then centrifuged again at for 1 min. (F) The supernatant (approximately 650 μ l) was removed to a new centrifuge tube and combined with 650 μ l of solution C4. (G) The resulting lysate preparation was vortexed vigorously for 5 s. (H) Next, 650 μ l of the vortexed solution were added to the kit's purification columns in new centrifuge tubes. (I) After sample binding to the columns, the columns were washed with 650 μ l of 100% ethanol followed by a wash with 500 μ l of Solution C5. (J) The columns were then dried by centrifugation (2 min at 13,000 rpm), prior to being transferred to a clean centrifuge tube.

- (4) *RNA Recovery and DNA Removal* –(A) RNA elution was performed with 50 μ l of Solution C6. (B) The purified RNA was treated (according to protocol provided by manufacturer) with AMBION

Table 2

Summary of demonstration variables and study samples.

	Cryogenic Preservation	Petroleum Liquids	<i>E. coli</i> Spike	Step 2 – Sample Pretreatment
S1A-1 to S1A-3 - Triplicate	No	Yes	Yes	Yes
S1A-4 to S1A-6 - Triplicate	No	Yes	Yes	No
S1B-1 to S1B-3 - Triplicate	No	No	Yes	Yes
S1B-4 to S1B-6 - Triplicate	No	No	Yes	No
S2-1 to S2-3 - Triplicate	Yes	Yes	No	Yes
S2-4 to S2-6 - Triplicate	Yes	Yes	No	No
S2-7 to S2-9 - Triplicate	Yes	Yes	No	Yes
S3A, S3B, S3C - Field triplicate*	Yes	Yes	No	Yes

* (each split for RNA and DNA characterization).

DNA-FreeTM DNase (Life Technologies, Grand Island, NY) to remove co-extracted DNA prior to (C) RNA quantification via optical density at 260 nm with a NanoDrop (ThermoScientific, USA).

- (5) *Reverse Transcription to cDNA* – RNA was reverse transcribed to cDNA utilizing the SuperScriptTM IV First-Strand Synthesis System (Invitrogen, Carlsbad, CA).
- (6) *Reverse transcription quantitative polymerase chain reaction (RT-qPCR)*– Assays were prepared as 25 μ l reactions in a 96-well plate as follows. (A) A master mix was prepared by adding 12.5 μ l of Power SYBR GreenTM qPCR reaction mix (2X) (Life Technologies, Grand Island, NY), 1.5 μ l forward and reverse primers (2.5 μ M), and 7.5 μ l PCR grade water. (B) Then, 23 μ l of the master mix was added to each reaction well, followed by (C) the addition of 2 μ l of 0.5 ng of cDNA template (based on OD₂₆₀) to each well. Commercially available genomic DNA (ATCC) was used as calibration standards. (D) Once the well plate was prepared, it was run using an ABI 7300 real-time PCR system (Applied Biosystems, Foster City, CA) programmed with the following thermocycling conditions: 95 °C for 10 min, followed by 40 cycles of 95 °C for 45 s, 56 °C for 30 s, and 60 °C for 30 s. (E) Dissociation curve analysis was conducted to confirm amplicon specificity. (F) For each target assay, primer sets and genomic DNA controls were chosen per current published methods including: (i) bacterial 16S rRNA transcripts [18], (ii) archaeal 16S rRNA transcripts [18], (iii) methanogens via *mcrA* [19] and *R15* [20]), (iv) anaerobic benzene degraders via *abcA* [21], (v) anaerobic alkane degraders via *assA*, primer set *assA2F/assA2R*, [22], and (vi) alkane oxidizers via *alkB* [23].
- (7) *PCR amplification* –(A) 15 to 20 μ l of cDNA and DNA from samples 3-A, B and C obtained from environmental triplicates were placed in centrifuge tubes and shipped on blue ice to Research and Testing Laboratory, LLC (Lubbock, TX) for analysis of active and total present microbial communities. (B) As a first step, and prior to sequencing the samples, Research and Testing Laboratory, LLC (Lubbock, TX) generated barcoded 16S rRNA amplicons via PCR amplification from the provided cDNA and DNA.
- (8) *Next Generation Sequencing* –(A) If generated, 16S rRNA amplicons (see step 7) were sequenced via Illumina MiSeq following methods described in [24]. (B) Sequence data analysis was performed by Research and Testing Laboratory, LLC using the RDP classifier in conjunction with the Silva database for taxonomic identification of the 16S rRNA sequences (<https://www.arb-silva.de/>). (C) Relative abundance (%) data generated by sequencing were used to construct histograms that present active and total microbial community composition for each sample, based on their phylogeny. Analyses were done at the genus level. In cases where genera were unclassified, higher level taxonomic identifications are reported, but grouped taxa shared >97% sequence identity.

Method demonstrations

Demonstration variables and details regarding core samples are presented in Table 2. Demonstration variables include (1) sample preservation, (2) *E. coli* spikes (3) absence/presence of petroleum liquids, and 4) inclusion/exclusion of pretreatment Step 2. Non-cryogenically preserved Site

1 samples, spiked with *E. coli*, provided a basis for resolving RNA recovery given a known amount of RNA. Due to the non-cryogenic preservation of Site 1 samples, little, if any, RNA was likely present in the sample prior to *E. coli* spiking. Cryogenically preserved petroleum impacted samples from Site 2, analyzed with and without Step 2, provided a basis for demonstrating the merits of our methods with respect to the quality of the recovered RNA. Site 3 data document an NSZD driven shift in microbial ecology by comparing communities observed at the RNA and DNA levels at the leading edge of a petroleum body; the plume leading edge was identified by additional, unpublished results from site investigations.

The following describes methods associated with samples from each of the study sites.

Site 1 - (A) *E. coli* was grown overnight in LB medium at 37 °C. (B) When the *E. coli* culture contained approximately 8×10^8 cells per ml, 3 ml (in two 1.5 ml aliquots) of culture were pelleted by centrifugation at 3000 rpm for 5 min. The supernatant was removed, and the pellet was re-suspended in 200 μ l of lysis buffer from the Powerlyzer™ DNA extraction kit (Qiagen, Germantown MD). S1A and S1B samples (30 g) were each homogenized with a mortar and pestle. (C) After homogenization, the S1A and S1B samples were subdivided into six subsamples (5 g each), and each subsample was spiked with approximately 2.4×10^9 *E. coli* cells suspended in 200 μ l lysis buffer. (D) Lastly, the spiked soil samples were flash frozen using liquid nitrogen to simulate cryogenic sample preservation and stored at -80 °C until RNA extraction. Samples were extracted following described methods including and excluding "Step 2", and after DNase treatment, RNA yielded by all samples was quantified as described in Step 4 and reported as μ g of RNA/ g of soil.

Site 2 - (A) RNA was extracted from triplicate subsamples of S2 (S2-1 to S2-3) and analyzed using the steps in our procedure. Similarly, triplicate subsamples of S2 (S2-4 to S2-6) were extracted excluding Step 2 (sample pre-treatment). (B) Four different masses of RNA (0.5, 1, 2.5 and 5 ng) were analyzed via RT-qPCR for each extraction to determine quantities of 16S rRNA transcripts present, as a function of ng of RNA. The objective of this experiment was to determine the suitability of the RNA obtained with or without the pretreatment (Step 2), for quantification of 16S rRNA via RT-qPCR. At a later time, other S2 samples were extracted in triplicate (S2-7 to S2-9) and analyzed via RT-qPCR, targeting common hydrocarbon degradation biomarker genes to verify that the RNA yielded by extraction with the newly developed method also was suitable for performing analyses relevant to degradation of petroleum liquids.

Site 3 - (A) RNA extracted from S3 field triplicates was analyzed via next-generation sequencing as detailed in Step 8. (B) In addition, DNA was also extracted from samples S3 A, B and C following methods described in [24] and sequenced following methods outlined in Step 8. (C) DNA and cDNA sequence data were used to identify microbial communities present, and active, in the surveyed soils. Objectives for sequencing DNA and cDNA obtained from field triplicate S3 samples included: (1) to evaluate the feasibility of performing sequencing analyses with cDNA obtained from RNA purified with our described method, and (2) to document additional information gained from performing sequencing analysis targeting microbes that are both present and active, instead of just microbes present.

Next-generation sequencing for S3 samples provided phylogenetic identities and % abundance for the active and present microorganisms. Sequence data were analyzed as described previously [24]. Briefly, The RDP classifier was used in conjunction with the Silva database for taxonomic placement of the 16S rRNA sequences analyzed (<https://www.arb-silva.de/>). Relative abundance (%) data were used to construct bar charts that show microbial community composition for each sample, based on their phylogeny.

Results

Fig. 2 shows RNA recovery from the *E. coli* spiked Site 1 soils with and without petroleum liquids and with and without Step 2. With Step 2, mean RNA recovery from petroleum impacted samples is 35% greater than without Step 2. By contrast, mean RNA recovery from samples without petroleum is 20% less with Step 2 than without it. However, results from Tukey adjusted p-value analysis performed on RNA yield after ANOVA showed that there were no significant differences in RNA recovery between samples extracted with or without Step 2, for both samples with and without petroleum liquids. A

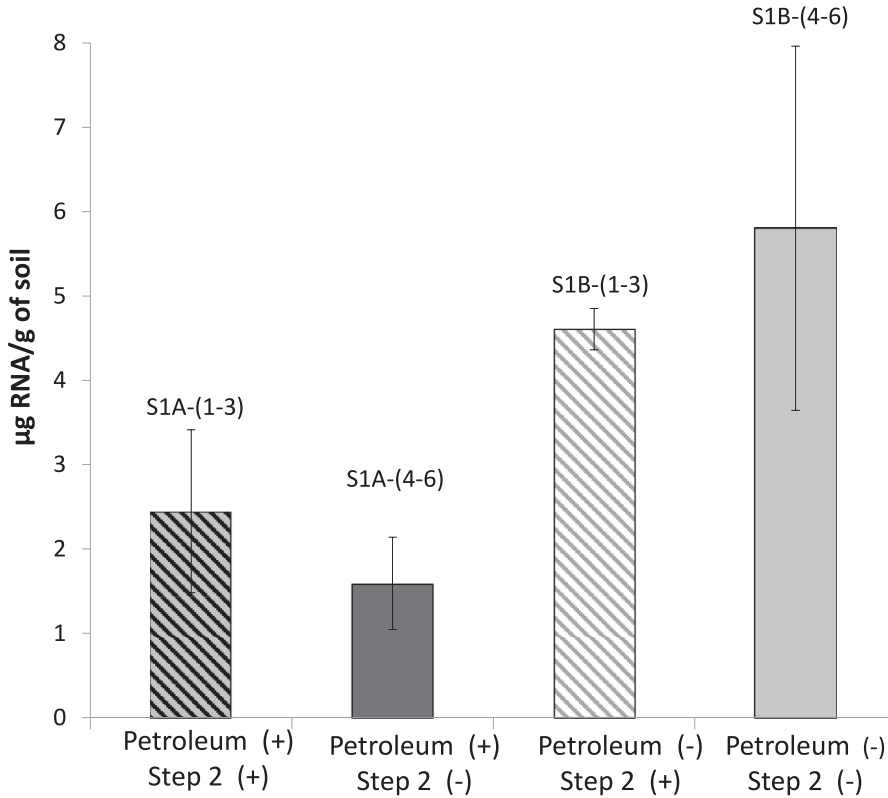


Fig. 2. Site 1 RNA yield per gram of soil. Error bars show 95% confidence intervals for an $n = 3$.

limitation of the Fig. 2 data is that the form and distribution of RNA in the *E. coli* spiked samples may not be representative of RNA in field soils. Microbes, and their RNA, in field soils occur primarily as thin biofilms in water bound to soil surfaces [25].

Step 2's main function is removing PCR inhibitors that compromise RNA utility for downstream analyses, including characterization of microbial ecology via transcript quantification with RT-qPCR or sequencing [26]. For example, the presences of metallic ions such as iron in solution [27] and organic acids are known to inhibit PCR reactions [28]. Both ferrous iron and organic acids are common byproducts of petroleum biodegradation [29], and thus likely to be present in petroleum impacted samples. Molecular assay disruption can include inaccurate quantification of transcripts or even non-detection of molecular targets. RNA was extracted from Site 2 soils with and without Step 2 and copied to cDNA. Results are reported as $\mu\text{g RNA/g}$ of soil.

Fig. 3 shows bacterial 16S rRNA transcripts as a function of the RNA mass used for samples from S2-1 to S2-6. Results of RT-qPCR analyses using Step 2 show a linear relationship between 16S rRNA transcripts quantified and RNA mass used for analysis, which demonstrates inhibitors were well removed by Step 2. In contrast, results of RT-qPCR excluding Step 2 do not show an appropriate linear relationship between 16S rRNA transcripts quantified and RNA mass. The lack of an appropriate linear relationships between transcript or gene quantities and template mass is attributable to inhibitors [30].

Continuing with Site 2, Fig. 4 plots *alkB*, *R15* and *mcrA* transcripts as a function of the mass of RNA analyzed for samples S2-7 through S2-9. RNA from samples S2-7 through S2-9 was extracted including pretreatment Step 2. RT-qPCR data shown on Fig. 4, targeting petroleum degradation processes,

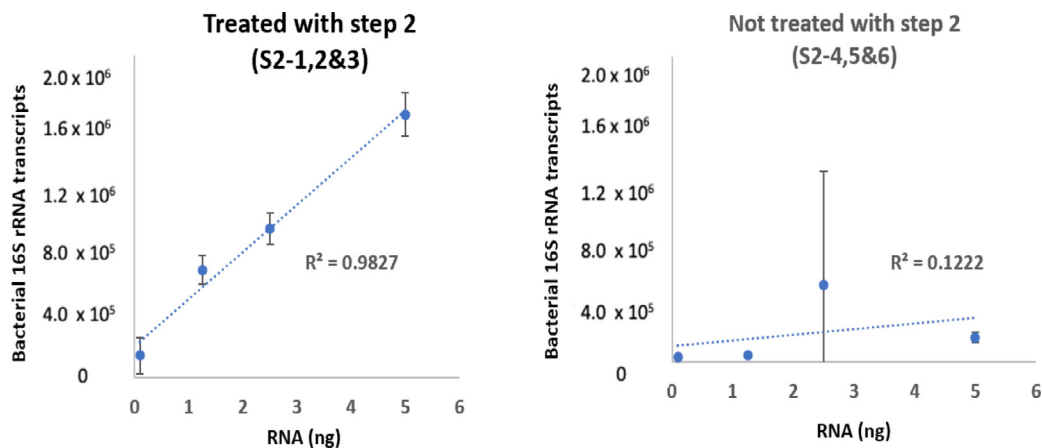


Fig. 3. Bacterial 16S rRNA transcripts versus RNA mass.

show a linear relationship between detected transcripts and template mass. Detected hydrocarbon degradation biomarkers in S2-7 through S2-9 include expressed *alkB* genes, which encode part of an alkane hydrolase. *alkB* is an enzyme involved in alkane degradation under aerobic conditions [31]. Detected methanogenic markers included *R15* (phylogenetic marker for methanogens commonly found in microaerophilic environments such as root zones) and *mcrA* (functional marker for methyl coenzyme M reductase). Additional petroleum degradation targets were assayed and not detected, including transcripts for *assA* (which encodes the α subunit of the first enzyme in the anaerobic alkane degradation pathway) and *abcA* (functional marker for anaerobic benzene degradation). Fig. 4 results illustrate that RNA obtained with our method is suitable for analysis of a variety of transcripts encoding hydrocarbon degradation enzymes.

Lastly, Site 3 soils were used to compare microbial communities based on RNA and DNA analyses. Fig. 5 shows the relative abundance of microorganisms at the identified grouped taxa level (grouped sequences shared >97% identity). Bacterial community analyses were generated by sequencing A) 16S rRNA bacterial genes (DNA) or B) 16S rRNA bacterial transcripts (RNA). Step 2 modifications were only employed for RNA. Microbial communities illuminated by RNA and DNA were distinct. At the DNA level, the community characterization shows taxa previously associated with anaerobic hydrocarbon degradation including putative fermenters such as: *Pelotomaculum* [32], unclassified *Clostridia* [33], Deltaproteobacteria [34], *Firmicutes* [35], unclassified *Anaerolineaceae* [36], and *Smithella* [35]. At the RNA level, Step 2 was employed to obtain RNA in support of resolving microorganisms that were active at the time of sampling. Interestingly, the active bacterial community was predominantly aerobic. Specifically, identified aerobic microorganisms previously associated with hydrocarbon degradation include *Citrobacter* [35], *Corynebacterium* [37], *Kluyvera* [38], and *Staphylococcus* [39].

Identifying an active aerobic microbial community in the analyzed soil is consistent with on-going historical disappearance of petroleum liquids and degradation products in groundwater in wells for the study area. Based on other site characterization efforts, the study area is known to lie at the leading edge of a large historical subsurface body of petroleum liquids. The sample analyzed contained coarse sediments (i.e., sands and gravel) and was collected from a subsurface zone composed of high-transmissivity sands and gravel beds (high-flow) that are interbedded with lower transmissivity silts (low-flow). At the sample collection site, large groundwater fluctuations (3–7 m) are driven by an adjacent river that is hydraulically connected to the hydrocarbon impacted aquifer. Riparian influences likely include the delivery of dissolved oxygen to the transmissive sand and gravel layers of the aquifer consistent with the observed active aerobic microbial populations (RNA) reported herein. Detailed chemical analysis of site samples completed via Fourier-transform ion cyclotron resonance

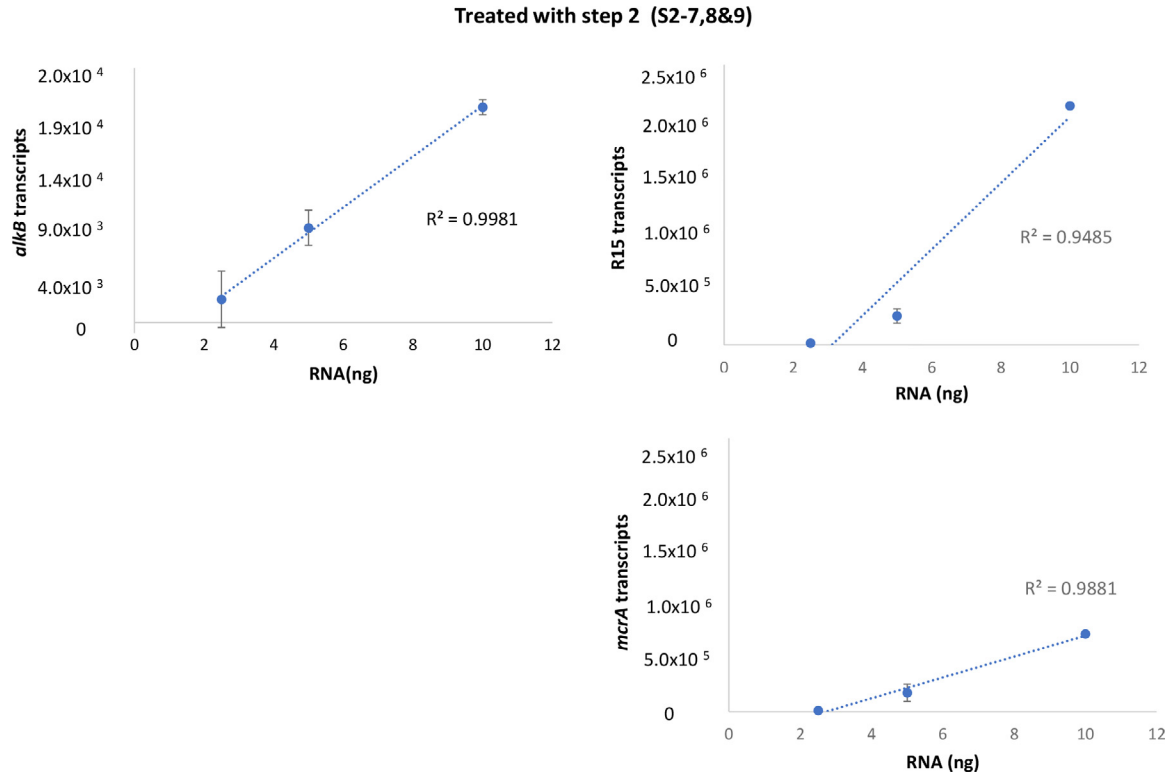


Fig. 4. *alkB*, R15 and *mcrA* transcripts versus RNA mass for sample triplicates S2 -7,8, and 9 extracted including treatment Step 2.

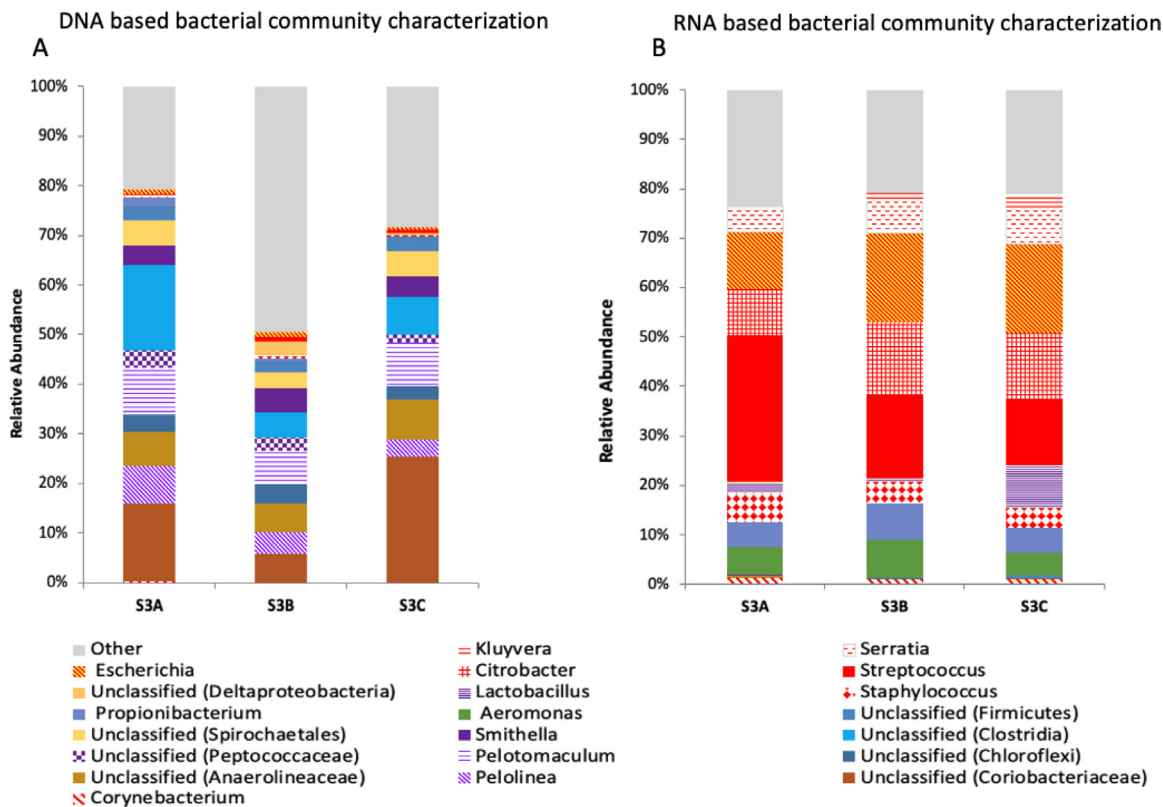


Fig. 5. Relative abundance of microorganisms based on DNA (A) and RNA (B) acquired from splits of cryogenic samples obtained from 9.3 m bgs from three adjacent borings. Analyses were done at the genus level. In cases where genera were unclassified, higher level taxonomic identifications are reported, but grouped taxa shared >97% sequence identity.

mass spectrometry further supports observed patterns in microbial ecology [40]. Also noteworthy in Fig. 5 are the similarities of identified communities in the three closely located collected cryogenic cores for both the DNA and RNA-based methods, suggesting methods are generally reproducible.

Conclusions

The pretreatment wash (Step 2) did not improve RNA yield but had no adverse effects on RNA yield. Importantly, clear improvements in purity of the extracted RNA were achieved with Step 2. Benefits of improved RNA purity include a reduced potential for failed or inhibited amplification of molecular targets, which would lead to inaccurate quantification of targeted genes or transcripts and potentially inaccurate relative abundance data from sequencing. Overall cleanup of nucleic acids prior to analyses, such as RT-qPCR and sequencing, can improve the quality of RNA-based characterizations of microbial communities. Lastly, results illustrate that RNA obtained with our method is suitable for the identification and quantification of a variety of transcripts associated with hydrocarbon degradation.

An intriguing aspect of this work is the potential to use differences between RNA- and DNA-based characterization of microbial communities to illuminate microbial community activity and dynamics, which play critical roles in progress to site restoration. Herein, the difference between the overall microbial communities present, observed via DNA-based tools, and the active microbial communities, observed via RNA-based tools, is striking. Hydrocarbon degradation kinetics are generally known to be up to an order of magnitude faster under more oxidizing conditions (aerobic and nitrate reducing) versus more reduced conditions (sulfate reducing and methanogenic conditions) [36]. Thus, our observation that the active microbial communities at the time of sampling were aerobic is consistent with on-going historical disappearance of petroleum liquids and related compounds in groundwater in wells in the study area, as described in unpublished site investigation reports. Critically, this insight into the putative role of aerobic microbial metabolism in remediation at study Site 3 would have been overlooked based on DNA analysis alone given that the DNA-based microbial characterization identified a predominantly anaerobic fermenting community. Less than 2% of the microbial community identified based on the DNA analysis was characterized as aerobic. Overall, our methods for extracting RNA from petroleum impacted soils, in combination with DNA data, provide a promising new tool for resolving evolving metabolic processes and biochemistry at sites progressing toward restoration via natural and/or imposed remedies. Ongoing work is focused on further explorations of concurrent RNA/DNA based characterization of microbially mediated processes contributing to restoration of sites impacted by petroleum hydrocarbons and other contaminants of concern.

Acknowledgments

The authors would like to thank Mo Bio technical personnel for guidance on modifications to their PowerSoil Powerlyzer™ DNA purification protocol to be suitable for RNA recovery. We would also like to thank the University Consortium for Field-Focused Groundwater Research for funding. Finally, we would like to thank Dr. Jacqueline Chaparro and Dr. Paige Wilson for technical input regarding PCR inhibitors in environmental samples and existing RNA purification methods.

Declaration of Competing Interest

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.

References

- [1] G.M. van der Kraan, M. De Ridder, B.P. Lomans, G. Muyzer, Sampling and nucleic extraction procedures from oil reservoir samples, in: *Applied Microbiology and Molecular Biology in Oilfield Systems*, Springer, 2010, pp. 7–16.
- [2] S. Garg, C.J. Newell, P.R. Kulkarni, D.C. King, D.T. Adamson, M.I. Renno, T. Sale, Overview of natural source zone depletion: processes, controlling factors, and composition change, *Groundw. Monit. Remediat.* 37 (3) (2017) 62–81.

- [3] T. Sale, H. Hopkins, A. Kirkman, Managing Risks at LNAPL Sites, American Petroleum Institute, 2018 (Bulletin No. 18).
- [4] Evaluating LNAPL remedial technologies for achieving project goals, Interstate Technology & Regulatory Council, LNAPLs Team Washington, DC, 2009.
- [5] P.D. Lundegard, P.C. Johnson, Source zone natural attenuation at petroleum hydrocarbon spill sites—II: application to a former oil field, *Groundw. Monit. Remediat.* 26 (4) (2006) 93–106.
- [6] N.J. Sihota, K.U. Mayer, M.A. Toso, J.F. Atwater, Methane emissions and contaminant degradation rates at sites affected by accidental releases of denatured fuel-grade ethanol, *J. Contam. Hydrol.* 151 (2013) 1–15.
- [7] K. McCoy, J. Zimbron, T. Sale, M. Lyverse, Measurement of natural losses of LNAPL using CO₂ traps, *Groundwater* 53 (2015) 658–667.
- [8] K. Karimi Askarani, E.B. Stockwell, K.R. Piontek, T.C. Sale, Thermal monitoring of natural source zone depletion, *Groundw. Monit. Remediat.* 38 (3) (2018) 43–52.
- [9] K.K. Askarani, T.C. Sale, Thermal estimation of natural source zone depletion rates without background correction, *Water Res.* 169 (2020) 115245.
- [10] D.R. Lovley, S.J. Giovannoni, D.C. White, J.E. Champine, E. Phillips, Y.A. Gorby, S. Goodwin, *Geobacter metallireducens* gen. nov. sp. nov., a microorganism capable of coupling the complete oxidation of organic compounds to the reduction of iron and other metals, *Arch. Microbiol.* 159 (4) (1993) 336–344.
- [11] N.R. Zeman, M.I. Renno, M.R. Olson, L.P. Wilson, T.C. Sale, K. Susan, Temperature impacts on anaerobic biotransformation of LNAPL and concurrent shifts in microbial community structure, *Biodegradation* 25 (4) (2014) 569–585.
- [12] M. Irianni-Renno, D. Akhbari, M.R. Olson, A.P. Byrne, E. Lefevre, J. Zimbron, M. Lyverse, T.C. Sale, S.K. De Long, Comparison of bacterial and archaeal communities in depth-resolved zones in an LNAPL body, *Appl. Microbiol. Biotechnol.* 100 (7) (2016) 3347–3360.
- [13] S. Kiaalhosseini, R.L. Johnson, R.C. Rogers, M.I. Renno, M. Lyverse, T.C. Sale, Cryogenic core collection (C3) from unconsolidated subsurface media, *Groundw. Monit. Remediat.* 36 (4) (2016) 41–49.
- [14] M. Olson, S. De Long, M. Irianni-Renno, W. Clayton, T. Sale, R. Johnson, Evaluating Long-Term Impacts of Soil-Mixing Source-Zone Treatment using Cryogenic Core Collection, Trihydro Corporation Laramie United States, 2017.
- [15] M.A. Zaitoun, C.T. Lin, Chelating behavior between metal ions and EDTA in sol-gel matrix, *The Journal of Physical Chemistry B* 101 (10) (1997) 1857–1860.
- [16] X. Sui, G.D. Ji, Impact of ultrasonic power density on elution of super heavy oil and its biomarkers from aging soils using Triton X-100 micellar solution, *J. Hazard. Mater.* 176 (1–3) (2010) 473–480.
- [17] Y.D. Livney, Milk proteins as vehicles for bioactives, *Curr. Opin. Colloid Interface Sci.* 15 (1–2) (2010) 73–83.
- [18] M.T. Suzuki, L.T. Taylor, E.F. DeLong, Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays, *Appl. Environ. Microbiol.* 66 (11) (2000) 4605–4614.
- [19] H. Juottonen, P.E. Galand, K. Yrjölä, Detection of methanogenic Archaea in peat: comparison of PCR primers targeting the *mcrA* gene, *Res. Microbiol.* 157 (10) (2006) 914–921.
- [20] T.J. McGenity, K.N. Timmis, B.N. Fernández, *Hydrocarbon and Lipid Microbiology Protocols*, Springer, 2016.
- [21] M.J. van der Waals, S. Atashgahi, U.N. Da Rocha, B.M. van der Zaan, H. Smidt, J. Gerritse, Benzene degradation in a denitrifying biofilm reactor: activity and microbial community composition, *Appl. Microbiol. Biotechnol.* 101 (12) (2017) 5175–5188.
- [22] C. Aitken, D. Jones, M. Maguire, N. Gray, A. Sherry, B. Bowler, A. Ditchfield, S. Larter, I. Head, Evidence that crude oil alkane activation proceeds by different mechanisms under sulfate-reducing and methanogenic conditions, *Geochim. Cosmochim. Acta* 109 (2013) 162–174.
- [23] G. Guo, F. Tian, K. Ding, L. Wang, T. Liu, F. Yang, Effect of a bacterial consortium on the degradation of polycyclic aromatic hydrocarbons and bacterial community composition in Chinese soils, *Int. Biodeterior. Biodegrad.* 123 (2017) 56–62.
- [24] M. Irianni-Renno, D. Akhbari, M.R. Olson, A.P. Byrne, E. Lefèvre, J. Zimbron, M. Lyverse, T.C. Sale, K. Susan, Comparison of bacterial and archaeal communities in depth-resolved zones in an LNAPL body, *Appl. Microbiol. Biotechnol.* 100 (7) (2016) 3347–3360.
- [25] D. Or, B.F. Smets, J. Wraith, A. Dechesne, S. Friedman, Physical constraints affecting bacterial habitats and activity in unsaturated porous media—a review, *Adv. Water Resour.* 30 (6–7) (2007) 1505–1527.
- [26] L.S. Toni, A.M. Garcia, D.A. Jeffrey, X. Jiang, B.L. Stauffer, S.D. Miyamoto, C.C. Sucharov, Optimization of phenol-chloroform RNA extraction, *MethodsX* 5 (2018) 599–608.
- [27] M. Sidstedt, P. Rådström, J. Hedman, PCR inhibition in qPCR, dPCR and MPS—mechanisms and solutions, *Anal. Bioanal. Chem.* 412(9) (2020) 2009–2023.
- [28] C.N. Albers, A. Jensen, J. Bælum, C.S. Jacobsen, Inhibition of DNA polymerases used in Q-PCR by structurally different soil-derived humic substances, *Geomicrobiol. J.* 30 (8) (2013) 675–681.
- [29] J.D. Coates, D.J. Lonergan, E.J.P. Phillips, H. Jenter, D.R. Lovley, *Desulfuromonas palmitatis* sp. nov., a marine dissimilatory Fe III reducer that can oxidize long-chain fatty acids, *Arch. Microbiol.* 164 (6) (1995) 406–413.
- [30] E.J. Kontanis, F.A. Reed, Evaluation of real-time PCR amplification efficiencies to detect PCR inhibitors, *J. Forensic Sci.* 51 (4) (2006) 795–804.
- [31] Z. Shao, W. Wang, Enzymes and genes involved in aerobic alkane degradation, *Front. Microbiol.* 4 (2013) 116.
- [32] K.E. Scherr, T. Lundaa, V. Klose, G. Bochmann, A.P. Loibner, Changes in bacterial communities from anaerobic digesters during petroleum hydrocarbon degradation, *J. Biotechnol.* 157 (4) (2012) 564–572.
- [33] N. Jiménez, H.H. Richnow, C. Vogt, T. Treude, M. Krüger, Methanogenic hydrocarbon degradation: evidence from field and laboratory studies, *J. Mol. Microbiol. Biotechnol.* 26 (1–3) (2016) 227–242.
- [34] D. Moreira, F. Rodríguez-Valera, P. López-García, Metagenomic analysis of mesopelagic Antarctic plankton reveals a novel deltaproteobacterial group, *Microbiology* 152 (2) (2006) 505–517.
- [35] D. An, D. Brown, I. Chatterjee, X. Dong, E. Ramos-Padron, S. Wilson, S. Bordenave, S.M. Caffrey, L.M. Gieg, C.W. Sensen, Microbial community and potential functional gene diversity involved in anaerobic hydrocarbon degradation and methanogenesis in an oil sands tailings pond, *Genome* 56 (10) (2013) 612–618.
- [36] B. Liang, L.-Y. Wang, S.M. Mbandinga, J.-F. Liu, S.-Z. Yang, J.-D. Gu, B.-Z. Mu, Anaerolineaceae and Methanosaeta turned to

- be the dominant microorganisms in alkanes-dependent methanogenic culture after long-term of incubation, *AMB Express* 5 (1) (2015) 37.
- [37] A. Linda, A. Bouziane, Petroleum-oil biodegradation by *Corynebacterium aquaticum* and *Pseudomonas aeruginosa* strains isolated from the industrial rejection of the refinery of ARZEW-Algeria, *World Appl. Sci* 18 (2012) 1119–1123.
- [38] S.K. Shekhar, J. Godheja, D. Modi, Hydrocarbon bioremediation efficiency by five indigenous bacterial strains isolated from contaminated soils, *Int. J. Curr. Microbiol. Appl. Sci.* 4 (3) (2015) 892–905.
- [39] C. Bhuvaneswar, G. Swathi, B. Bhaskar, T. Munichandrababu, W. Rajendra, Effective synergetic biodegradation of diesel oil by bacteria, *Int. J. Environ. Biol.* 2 (4) (2012) 195–199.
- [40] O. Bojan, M. Irianni-Renno, A. Hanson, C. Huan, R. Young, S.K. De Long, T. Borch, T.C. Sale, A.M. McKenna, J. Blotevogel, Discovery of oxygenated hydrocarbon biodegradation products at a late-stage petroleum release site, *Energy Fuels*, In review.