Diversity and Similarity of Microbial Communities in Petroleum Crude Oils Produced in Asia

Kunio YAMANE,1,† Hideaki MAKI,2 Tsuyoshi NAKAYAMA,3 Toshiaki NAKAJIMA,3 Nobuhiko NOMURA,3 Hiroo UCHIYAMA,3 and Motomitsu KITAOKA1

1National Food Research Institute, 2-1-12 Kannondai, Tsukuba-shi, Ibaraki 305-8642, Japan
2National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba-shi, Ibaraki 305-8506, Japan
3Graduate School of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba-shi, Ibaraki 305-8572, Japan

Received April 8, 2008; Accepted July 24, 2008; Online Publication, November 7, 2008
[doi:10.1271/bbb.80227]

To understand microbial communities in petroleum crude oils, we precipitated DNA using high concentrations of 2,2,4-trimethylpentane (isooctane) and purified. Samples of DNA from five crude oils, (Middle East, 3; China, 1; and Japan, 1) were characterized based upon their 16S rRNA gene sequences after PCR amplification and the construction of clone libraries. We detected 48 eubacterial species, one cyanobacterium, and one archaeon in total. The microbial constituents were diverse in the DNA samples. Most of the bacteria affiliated with the sequences of the three oils from the Middle East comprised similar mesophilic species. Acinetobacter, Propionibacterium, Sphingobium and a Bacillales were common. In contrast, the bacterial communities in Japanese and Chinese samples were unique. Thermophilic Petrotoga-like bacteria (11%) and several anaerobic-thermophilic Clostridia- and Synergistetes-like bacteria (20%) were detected in the Chinese sample. Different thermophiles (12%) and Clostridia (2%) were detected in the Japanese sample.

Key words: petroleum crude oil; alkane composition; microbial community; 16S rRNA gene diversity; precipitation with isooctane

Petroleum crude oils are thought to be formed by a complex and incompletely understood series of chemical reactions from organic materials laid down in previous geological eras. Many chemical compounds in crude oils have been fractionated and identified using advanced analytical methods, especially capillary GC-MS.1 In contrast, knowledge of the biological species and microbial communities that are associated with the generation of crude oils has remained limited.2,3 Crude oil reservoirs are generally found in deep geological formations characterized by high temperatures, high pressures and an anaerobic environment. Culture-based methods have been extensively implemented to improve understanding of petroleum microbiology. Nevertheless, little is understood about the microbial diversity and communities of these ecosystems, since many of the microorganisms harboring under these conditions are non-culturable using present experimental techniques. Molecular methods based on PCR have proven effective at characterizing complex microbial assemblages in environmental samples. In particular, sequence analyses of 16S rRNA genes in samples have revealed previously unrecognized microbial diversity in various habitats, even in the terrestrial subsurface.4–6)

The application of culture-independent techniques to petroleum microbiology contributes to a more precise characterization of microbial communities in subsurface oil reservoirs. Dot-blot hybridization with functional gene probes and 16S rRNA gene sequence analysis have been applied to identify the sulfate-reducing bacterial population inhabiting a low temperature, water-flooded well in Western Canada.7,8) Both 16S rRNA gene sequence phylogenetic analysis and enrichment culture techniques have been used to characterize thermophilic microbial assemblages in the Miocene Monterey Formation, a prominent high-temperature, oil-bearing formation in California.9) In addition, the bacterial communities in high-temperature, water-flooded petroleum reservoirs of continental (Huabei) and offshore (Qinghuang) oilfields in China have been investigated using the 16S rRNA gene approach.10,11) Samples of DNA for these studies were prepared from oilfield production water, and the aqueous fraction was separated from oil-water emulsions. To identify the microbial communities in crude oils, we attempted to prepare

† To whom correspondence should be addressed. Tel: +81-29-838-8071; Fax: +81-29-838-7321; E-mail: yamanek@affrc.go.jp

Abbreviations: GC-MS, gas chromatography coupled with mass spectrometry; TLC, thin-layer chromatography; FID, flame ionization detector; DGGE, denaturing gradient gel electrophoresis
These samples were stored at 4°C oil field (Teikoku Oil, Niigata, Japan) in June, 2005. Sampled directly from a wellhead at the Minami-Aga Oil Co., Ltd. (Chiba, Japan), were provided by them in October, 2006. Minami-Aga crude oil (MAG) was used to show the diversity and similarity of eubacterial and archaeal communities in five crude oils from different geographic regions, based on 16S rRNA gene sequences.

Materials and Methods

Crude oil samples. Arabian extra light (AEXL), Arabian medium (ARM), and Qatar land light (QTL) crude oil samples from the Middle East, and Daqing crude oil (TK) imported from China, which had been stored in central research laboratories, Idemitsu Kosan Oil Co., Ltd. (Chiba, Japan), were provided by them in October, 2006. Minami-Aga crude oil (MAG) was sampled directly from a wellhead at the Minami-Aga oil field (Teikoku Oil, Niigata, Japan) in June, 2005. These samples were stored at 4°C until analysis.

Asphaltene content in crude oils. Asphaltene contents in crude oils were measured by TLC coupled with FID, (Iatroscan MK-5, Iatron, Tokyo). Oil components were separated by TLC using three solvent systems, and the amounts of asphaltene (%) remaining at the origins of TLC were calculated after FID detection.

GC-MS analysis of n-alkane components in crude oils. The composition of n-alkanes in the five crude oils was analyzed by GC-MS (QP5050, Shimadzu, Kyoto, Japan) using a 26 m × 0.25 mm i.d. (0.25 μl film) DB-5 fused silica capillary column (J & W Scientific, Folsom, CA) with helium (2.0 ml/min) as the carrier gas. The injection and detector temperatures were 320°C. Separation proceeded at the following temperatures: 5 min hold at 50°C, ramp to 320°C at 5.5°C/min, and 3 min hold at 320°C.

Preparation of DNA from crude oil samples. After removing volatile components from the crude oils (except for the TK sample), 25 ml of each oil was vortex-mixed with an equal volume of isooctane (2,2,4-trimethylpentane) and left overnight at 4°C. The samples were thoroughly mixed and sedimented by centrifugation at 5,000 × g for 60 min at 4°C. The precipitates were suspended in 50 ml of isooctane, and sedimented once again by centrifugation at 5,000 × g for 30 min at 4°C. They were washed twice with 10 ml of isooctane and dried in vacuo. Then DNA was extracted 3 times from each dried sample (29–200 mg) using the UltraClean™ soil DNA isolation kit (MO BIO Laboratories, Carlsbad, CA) according to the manufacturer’s instructions. Each sample yielded about a 50 μl of DNA solution. The wax content (17.1%, w/w) in the TK sample was higher than that in the others (2–3%, w/w). The sample assumed a solid form at room temperature. We melted the TK sample at 55°C and thoroughly mixed it with 3 volumes of isooctane at 55°C. The mixture was placed at room temperature overnight and centrifuged and then the precipitate was washed at room temperature as described above. The washed precipitate was dried, and 50 μl of DNA solution was extracted using the DNA isolation kit. As a control DNA sample, Blank DNA was prepared from 25 ml of isooctane in the same manner instead of crude oil.

To verify that the DNAs were efficiently extracted from crude oils, Pseudomonas aeruginosa cells were introduced with mixing into oils, and then DNA was prepared from them. P. aeruginosa cells cultured in LB medium overnight were precipitated twice in acetone and resuspended in absolute acetone. Then we introduced with mixing 0, 0.4, 1.0, and 2.0 × 10^7 of the cells per ml of QTL oil samples. After 1 week at room temperature, the mixtures were precipitated with 50% isooctane and then washed and dried as described above. We then extracted 30 μl of DNA solution from 10 ml of QTL oil using the soil DNA Isolation Kit (MO BIO Laboratories). The control comprised 30 μl of DNA solution extracted from 1.0 × 10^7 acetone-dried P. aeruginosa cells using the same kit.

Determination of DNA content. The DNA contents in the samples extracted from crude oils were determined by the absorbance at 260 nm using a NanoDrop ND1000 spectrometer (NanoDrop Technology, Wilmington, DE), and by real-time PCR using primers 357F (5'-CCTACGGAGGCGACCAG-3') and 518R (5'-GTATTACCGCGGCTGCTG-3'), which can detect a wide range of eubacterial 16S rRNA genes. The DNA concentration of euBacteria was determined in 20 μl reaction volumes containing 1 × Power SYBR® Green PCR Master Mix (Applied Biosystems, Foster City, CA) on a 7300 real-time PCR platform (Applied Biosystems) under the following conditions: initial denaturation at 95°C for 10 min and 40 cycles of 30 sec at 94°C, 60 sec at 58°C, and 30 sec at 72°C. Standard curves for real-time PCR were generated using chromosomal DNA from P. aeruginosa. The curves for a triplet series containing a dynamic range of 4 orders of magnitude (9 ng/μl to 0.009 ng/μl) closely agreed.

Amplification of eubacterial and archaeal 16S rRNA gene sequences. We amplified the 16S rRNA genes in the bulk DNA extracted from crude oil samples and any DNA in the Blank by PCR using Nova Taq Hot start DNA polymerase (Novagen, Madison, WI), 1 × AmpDirect Plus buffer (Shimadzu, Kyoto, Japan), and the bacteria-specific primer set, 63F (5'-CAGGCCCTAACA-
CATGAAGTC-3') and 1387R (5'-GGGGCGGWGTGTA-CAAGGC-3') \(^{27}\) which can amplify approximately 1350 bp fragments. Amplification proceeded in duplicate tubes containing 50 µL of reaction mixtures on a GeneAmp PCR system 9700 (Applied Biosystems). The thermal cycler program included initial denaturation at 95°C for 10 min, 40 cycles of 94°C for 30 sec, 58°C for 60 sec, and 72°C for 90 sec and a final elongation step 72°C for 10 min. The PCR products were purified using the UltraClean PCR Clean-up DNA purification kit (MO BIO Laboratories, Carlsbad, CA), and dissolved in 30-µL volumes. The purified PCR products were cloned into TOPO-TA plasmid using the TOPO-TA cloning kit (Invitrogen, Carlsbad, CA), and then transferred into Escherichia coli DH5α cells (Takara Bio, Otsu, Japan) to construct 16S rRNA gene libraries.

Archaeal 16S rRNA genes were amplified using nested PCR \(^{18}\) and primer sets, 109F (5'-AMDGCTCA-GTAACAGCT-3') and 1490R (5'-GGHTACCTTTGTT-ACGACTT-3') for the first, and 344F (5'-ACCGGGG-YGCACGGCCGGA-3') and 912R (5'-CTCCCCC-GCCAATTCCTTTA-3') for the second round of amplification. The PCR products were purified using the UltraClean PCR Clean-up DNA purification kit, and libraries of DNA samples were constructed using the E. coli system as described above.

Plasmid DNA preparation and sequencing. Plasmid DNA in each E. coli transformant of the 16S rRNA gene libraries was prepared using the QIAprep 8 Tube Microprep kit (Qiagen, Valencia, CA), and served as sequencing templates. Sequences were determined using an automated ABI PRISM 310 genetic analyzer (Perkin Elmer, Foster City, CA). For the eubacterial 16S rRNA gene sequences, we sequenced 80 to 100 clones for each library using the 357F primer. Sequences with >97% identity were considered identical. Representative clones were then fully sequenced using the RV (5'-CAGGAAAAACAGCTATGACC-3') and M4 (5'-GTAGGAT-CGCCAGTGAGC-3') primers. \(^{19}\) Archaeal 16S rRNA gene sequences of about 600 bp were determined using the universal primer. RV. The sequences were checked for chimeras using the CHIMERA_CHECK function of the ribosomal data base Project II. \(^{20}\) The sequences were compared with similar sequences of reference organisms by BLAST search. \(^{21}\) The 16S rRNA gene sequences were submitted to DDBJ/EMBL/GenBank and have been assigned accession nos AB368989 to AB369063 and AB439294 to AB439299.

**Results**

Characteristics of crude oil samples

Figure 1 summarizes the key characteristics of the five crude oil samples. GC-MS chromatograms showed that the n-alkane components (C\(_{10}\) to C\(_{30}\)) in the AEXL, ARM, and QTL samples were almost identical (Fig. 1A–C), although relative viscosity (RV), asphaltene content (AS), and specific gravity (SG) differed among them. The major n-alkanes of these three samples consisted of C\(_{11}\) to C\(_{20}\) with a small proportion of longer n-alkanes (C\(_{21}\) to C\(_{30}\)). ARM oil sample showing high viscosity, showed a high asphaltene content. In contrast, signals (▼) of unknown hydrocarbons were found between C\(_{12}\) and C\(_{17}\) n-alkanes in the GC-MS chromatogram from the MAG oil sample (Fig. 1D). Furthermore, its relative viscosity and the asphaltene content were very low. The content of longer chain n-alkanes (C\(_{17}\) to C\(_{30}\) and relative viscosity of the TK oil were very high (Fig. 1E), and the sample included 17.1% (w/w) wax.

No bacterial or fungal colonies arose from the crude oil samples on agar plates containing a 1/10 concentration of Luria-Bertani (LB) medium and a 1/100 concentration of Nutrient Broth containing 100 µg/mL of yeast extract (Difco) as far as tested, indicating that the crude oils did not contain viable microbes capable of growth under the tested conditions.

Efficiency of DNA precipitation from crude oils using isooctane as the precipitant

To verify the effectiveness of our DNA extraction from crude oils, we introduced with mixing the indicated amounts of *P. aeruginosa* cells into QTL oil sample. As Fig. 2 indicates, DNA was efficiently extracted from both QTL oil sample alone and mixtures of crude oil and *P. aeruginosa* cells. The recovered DNA from the mixtures was in proportion to the added cells. The two methods detected a very small amount of DNA in crude oil alone. The low content of DNA determined by real-time PCR compared with that determined at A\(_{260}\) (less than 10%) in the sample from QTL oil alone is to be attributed to DNA contents derived from organisms other than eubacteria and to contamination by other materials also having absorbance at 260 nm in the crude oil. Therefore, as the figure suggests, DNA was almost quantitatively extracted. Based on the 16S rRNA gene DNA content determined by real-time PCR and the average copy number of 16S rRNA genes (3.6 copies in the bacterial genome), \(^{22}\) the recovered DNA content corresponded to 10\(^4\)–10\(^5\) cells/mL. This abundance is similar to the range of 2.5 × 10\(^3\) to 1.42 × 10\(^6\) cells/ml determined in petroleum reservoirs \(^9\) and in deep subsurface habitats by direct cell counting. \(^{23}\)

The PCR products of a DNA preparation extracted from QTL crude oil containing 0.4 × 10\(^7\) *P. aeruginosa* cells as the template and the primer set, 357F and 907R (5'-CCGTAATTCTTTGAGT-3'), \(^{16}\) were inserted into TOPO-TA vector. All sequences of five *E. coli* transformants selected at random were 100% identical to the sequence of the corresponding region of the *P. aeruginosa* 16S rRNA gene.

Analysis of eubacterial 16S rRNA gene libraries

Table 1 summarizes the closest cultured strains, sequence identity, and numbers of identical sequences in each relative strain. The sequences of each library
Fig. 1. Comparison of Relevant Physicochemical Characteristics of Petroleum Crude Oils, AEXL (Arabian extra light, Saudi Arabia), ARM (Arabian medium, Saudi Arabia), QTL (Qatar land light, Qatar), MAG (Minami-Aga, Niigata, Japan), and TK (Daqing, China).

- Total ion chromatogram (scan mode) of hydrocarbons was analyzed by GC-MS. Relative viscosity (RV) was determined against water using an Ostwald viscometer at 50°C. Asphaltene contents (AS) (%) were determined by TLC-FID. Specific gravity (SP) is represented as g/cm³ at room temperature. Numbers represent carbon atoms in n-alkane backbone. ▼ Signals of unknown hydrocarbons.

- RV: 3.7
- AS: 0.2%
- SG: 0.822

- RV: 10.8
- AS: 7.1%
- SG: 0.869

- RV: 3.8
- AS: 0.5%
- SG: 0.831

- RV: 2.1
- AS: 0.3%
- SG: 0.813

- RV: 26.4
- AS: 5.8%
- SG: 0.852

Fig. 2. DNA Concentration Measured by a A₂₆₀ Using NanoDrop ND1000 Spectrometer (A) and by Real-Time PCR Based on 16S rRNA Genes (B).

DNA was extracted from QTL crude oil alone and from the oil samples to which 0.4, 1.0, and 2.0 × 10⁹ Pseudomonas aeruginosa cells/ml had been mixed. Shaded bars show DNA content corresponding to crude oil alone, and open bars show DNA contents corresponding to known numbers of P. aeruginosa that had been mixed. Closed bars show DNA contents corresponding to 1 × 10⁷ cells that had been dried after suspended in acetone. At the bottom of the closed bars, the DNA contents corresponding to crude oil alone (*) are shown for comparison.
revealed a diverse eubacterial population. Ten to 16 phylotypes were detected in each. We detected 48 eubacterial species that were included among 25 genera and one cyanobacterium in total. Figure 3 shows distribution of sequence types from the eubacterial 16S rRNA gene libraries in bacterial groups.

The closest bacteria affiliated with the sequences of the AEXL, ARM, and QTL libraries were similar to each other except for the high numbers of a cyanobacterium in the ARM library. *Acinetobacter*, *Propionibacterium*, *Sphingobium*, and a *Bacillales* bacterium were common. In addition, the *Corynebacterium*, *Methylobacterium*, and *Staphylococcus* genera were also detected in the DNA libraries. They were made up of alphaproteobacteria, gammaproteobacteria, actinobacteria, and firmicutes. All the analyzed sequences had high identity (more than 97%) with previously reported sequences, except for that of the Bacillales bacterium Gsoil 1105 (AB245375). Cloned sequence having 87% identity to Gsoil 1105 had 94% identity to that of

---

**Table 1.** Closest Cultured Strains of Eubacteria, Cyanobacterium, and Archaeon Based on 16S rRNA Gene Libraries from DNA of AEXL, ARM, QTL, MAG, and TK Crude Oils, and from BLANK DNA
an uncultured Firmicutes bacterium, MoB-G6-114 (EF016847), the DNA of which was prepared from hyper-arid Atacama desert soil.

In contrast, the bacteria detected from the MAG and TK libraries were rich in unique strains. A half of the cloned sequences were similar to those of Corynebacterium xerosis and an uncultured Actinobacterium. Many analyzed sequences had low identity with reference organisms. The novel thermophilic bacteria Petrotoga mobilis and P. halophila, isolated from a producing oil well in the North Sea\(^{24}\) and from an off-shore producing oil well in Congo\(^{25}\) respectively, were detected in the TK library. Furthermore, the closest cultured relatives of anaerobic and thermophilic Clostridia and Synergistetes (Anaerobaculum thermoterrnum, U50711, Desulfutibacter alkalitolerans, AY338171, Desulfomaculum thermosaporovar, Z26315, and Moorella thermoacetica, CP000232) were also detected, although the identity between the cloned sequences and that of closest relatives was low (Table 1). Figure 4 shows phylogenetic affiliation of the detected 16S rRNA gene sequences with relatives of Clostridia, Synergistetes, and Bacillales. The TK library was enriched with Actinobacteria, Thermotogae, and Firmicutes, whereas the Japanese MAG library was rich in Actinobacteria and Gammaproteobacteria. Of 58 tested clones from the MAG library, 23 clones (40%) were to be placed in Proteobacteria (Table 1). Gammaproteobacteria, nine clones were identified as Marinobacterium sp. IC961 (AB196257), a carboxal-degrading bacterium isolated from environment samples, and eight clones were distantly related to Methylohalobius crimeensis (AJ581837), a moderately halophilic, methanotrophic bacterium isolated from hypersaline lakes of Crimea. Five clones fell into Deltaproteobacteria and were identified as Desulfomicrobium thermophilum (AY464939), a sulfate-reducing thermophile isolated from a terrestrial hot spring in Colombia. One clone in Betaproteobacteria was distantly related to Petrobacter sp. DM-3 (DQ539621), a thermophilic oil-degrading bacterium

Fig. 3. Proportion of the Detected Sequence Types from the Eubacterial 16S rRNA Gene Libraries of AEXL, ARM, QTL, MAG, and TK Oil Samples.

Fig. 4. Phylogenetic Tree of the Clostridia, Synergistetes and Bacillales 16S rRNA Gene Phylotypes of Bacterial Sequences (PT-TK-B in bold) That Were Detected in the DNA Library from TK Crude Oil and Closely Related Sequences by the BLAST Search. The topology shown was constructed by the maximum likelihood method. Generalized time reversible (GTR) model was used with estimated proportion of invariant sites and rate heterogeneity among variable sites approximated as discrete gamma distribution.\(^{30}\) Numbers around internodes indicate bootstrap values (\(>50\%\)) in maximum parsimony, neighbor-joining, and maximum likelihood analyses (100/100/100 replications respectively) using PAUP* var. 4.0b10. Geobacter uraniireducens (EF527427) was used as the outgroup sequence.
isolated in China. One Firmicutes bacterium was closely related to Tindallia texcocoense (DQ234901), a haloalkaliphilic Clostridia isolated from Lake Texcoco, Mexico. Therefore, the detected bacteria in the MAG library were quite different from those from the TK library.

Agrobacterium tumefaciens, Mycobacterium flavescentis, Propionibacterium acnes, and Variovorax paradoxus were the closest relatives of DNA sequences found in the Blank DNA library. Propionibacterium acnes is a commensal bacterium of human skin, indicating that some of the Propionibacteria in the libraries of crude-oil DNA are contaminants that arise during DNA preparation. Some of the V. paradoxus found in ARM and TK libraries will are also contaminants.

16S rRNA gene libraries to detect archaea

A common sequence with 99% identity to a cultured archaeon, Methanoseta harundinacea (AY970347), a novel acetate-scavenging methanogen, appeared at high frequency in the MAG, QTL, and TK libraries. The sequences obtained from three libraries were identical and also similar (99%) to that of an uncultured Methanosetaeaceae archaeon clone BR34ARC_A09 (AY324377), a methane-oxidizing archaean that belongs to the genus Methanoscinales, but we did not detect any sequence homology to the archaean 16S rRNA in the AEXL, ARM, or Blank DNA libraries.

Phylogenetic analysis of cloned sequences related to cyanobacteria

Figure 5 shows a phylogenetic tree that includes the cloned PT-ARM-B8 sequence and sequences of the 16S rRNA gene of cyanobacterial relatives. The PT-ARM-B8 sequence is representative of 13 sequences among 54 tested clones of the ARM library. Another 12 sequences had more than 99.5% identity with PT-ARM-B8. This PT-ARM-B8 sequence had 92% identity to that of a cultured strain, Chroococcidiopsis sp. CC3-2 (DQ914865), from a hyper-arid desert in China. The sequence had 97% identity with 16S rRNA sequences from the uncultured bacterium JSC2-G9 (DQ532182) and the uncultured cyanobacterium FAC34 (DQ451473), from which DNA was obtained at the Johnson Space Center and from a forest soil in Taiwan respectively. Furthermore, PT-ARM-B8 had 93% identity with the uncultured cyanobacterium HAVO mat 11 from a Cyanobacterial mat in a lava cave at Hawaii Volcanoes National Park (EF032781) and to the unicellular thermophilic cyanobacterium tBTRCCn 28 from thermal springs in Jordan (DQ471448). In contrast, the nearest authentic cyanobacterial strains, Microcoleus glaciei and Phormidium murrayi, belonging to Oscillatoriales, had 91% identity. The phylogenetic position of PT-ARM-B8 also differed completely from those of other typical cyanobacteria, such as Anabaena, Synechococcus, and Synechocystis. Consequently, the cyanobacterium having the PT-ARM-B8 sequence ap-
pears to be unique. Detection of PT-ARM-B8 homologs in the sample of ARM was reproducible.

Discussion

Schlippers et al.\(^{23}\) reported that a large fraction of sub-seafloor prokaryotes is alive, even in very old (16 million years) and deep (<400 meters) sediments, and that all detectable living cells are bacteria with turnover rates of 0.25 to 22 years. Bacterial cells have been directly counted in oils and in samples from deep subsurface habitats.\(^{9,23}\) The DNA concentration found in the QTL oil that was measured by real-time PCR supported this thesis. If such bacteria are present in crude oils, we assumed that they could be precipitated by centrifugation as complexes with crude oil ingredients using high isoctane concentrations. We extracted microbial DNA from isoctane precipitates of crude oils using a commercially available soil DNA extraction kit and then amplified the DNA by PCR. A measurable amount of DNA was extracted from QTL crude oil alone, and quantitative amounts of DNA were also prepared from crude oil samples containing known numbers of *P. aeruginosa* cells. Therefore, this simple and direct method of preparing DNA from crude oils is useful in PCR-based molecular methods to characterize microbial communities and their diversity in crude oil reservoirs and stockpiles.

Most of the bacteria detected from the three Middle Eastern oil samples were facultative to mesophilic and previously reported bacteria. The common and dominant bacteria, *Acinetobacter* and *Sphingobium*, in the three samples are often strictly aerobic organisms and have been well known for oil-degrading capacity.\(^{27–29}\) It is possible that those bacteria thrived in the oils after recovery from reservoirs by utilizing oil components as C-sources, and became the dominant strains. In contrast, unique bacteria were detected in the oil samples from Japan and China. Oil-specific and thermophilic *Petrotoga*-like strains and thermophilic-anaerobic *Clostridia*- and *Synergistetes*-like strains were detected in the DNA of the TK oil. *Petrotoga*-like strains have been isolated from several high-temperature petroleum reservoirs worldwide,\(^{9–11,24,25}\) suggesting that these thermophiles are common components of specialized, hot geothermal subsurface environments. We also detected five phylotypes of the anaerobic bacteria *Clostridia* and *Synergistetes* (Fig. 4). The closest cultured strains were also thermophilic. The low identity between the detected and reported sequences in the closest cultured strains indicates that the bacteria with detected sequences are unique to this oil sample. The *Petrotoga*-, *Clostridium*-, and *Synergistetes*-like strains seemed to be adapted to thrive in deep subsurface reservoirs that are characterized by high-temperature, high-pressure and anaerobic systems. Consequently, they might be indigenous to petroleum reservoirs. The microbial community found in the MAG oil sample was also unique. We detected different types of thermophiles (sulfate-reducing *Desulfomicrobium* and oil-degrading *Petrobacter*) and a halophilic *Clostridium* (halophilic *Tindalia*). The identity between the cloned and reported sequences was also low. The diversity of the microbial communities detected in the DNA of crude oil samples indicated the heterogeneous nature of oils, and that our DNA preparation method using isoctane was useful to characterize the microbial communities of crude oils from reservoirs and stockpiles.

Crude oils are at high risk of contamination with microorganisms during drilling, transportation, and storage as well as during DNA extraction in laboratories. Since contaminating microorganisms cannot be distinguished from indigenous microorganisms, we believe that total DNA must first be analyzed, and then indigenous microorganisms can be selected by metagenome analysis.

The present study revealed significant diversity in the *n*-alkane composition of the five crude oils. The GC-MS profiles of the three oils from the Middle East were almost identical, whereas those of Japanese and Chinese oil samples were unique to each. The relationships between the *n*-alkane profiles and microbial diversity in crude oil samples remain unknown and await elucidation.

Acknowledgments

We thank Masato Kurihara (Idemitsu Kosan Co., Ltd.) and Tooru Nakao (Teikoku Oil Co., Ltd.) for providing crude oil samples, and Takashi Inaoka, National Food Research Institute, for technical assistance with real-time PCR procedures. We are grateful to N. Foster for critical reading of the manuscript.

References


