Massive dominance of *Epsilonproteobacteria* in formation waters from a Canadian oil sands reservoir containing severely biodegraded oil

Casey R. J. Hubert,†*a* Thomas B. P. Oldenburg,a Milovan Fustic,b,‡ Neil D. Gray,a Stephen R. Larter,b Kevin Penn,c,‡ Arlene K. Rowan,a,‡ Rekha Seshadri,a,‡ Angela Sherry,a,‡ Richard Swainsbury,a Gerrit Voordouw,a Johanna K. Voordouw,a and Ian M. Head1

1School of Civil Engineering and Geosciences, Newcastle University, Newcastle upon Tyne NE1 7RU, UK.
2Petroleum Reservoir Group, Department of Geoscience, University of Calgary, Calgary, Canada, T2N 1N4.
3J. Craig Venter Institute, San Diego, CA 92093, USA.
4Department of Biological Sciences, University of Calgary, Calgary, Canada, T2N 1N4.

**Summary**

The subsurface microbiology of an Athabasca oil sands reservoir in western Canada containing severely biodegraded oil was investigated by combining 16S rRNA gene- and polar lipid-based analyses of reservoir formation water with geochemical analyses of the crude oil and formation water. Biomass was filtered from formation water, DNA was extracted using two different methods, and 16S rRNA gene fragments were amplified with several different primer pairs prior to cloning and sequencing or community fingerprinting by denaturing gradient gel electrophoresis (DGGE). Similar results were obtained irrespective of the DNA extraction method or primers used. Archaeal libraries were dominated by *Methanomicrobiales* (410 of 414 total sequences formed a dominant phylotype affiliated with a *Methanoregula* sp.), consistent with the proposed dominant role of CO2-reducing methanogens in crude oil biodegradation. In two bacterial 16S rRNA clone libraries generated with different primer pairs, >99% and 100% of the sequences were affiliated with *Epsilonproteobacteria* (n = 382 and 72 total clones respectively). This massive dominance of *Epsilonproteobacteria* sequences was again obtained in a third library (99% of sequences; n = 96 clones) using a third universal bacterial primer pair (inosine-341f and 1492r). Sequencing of bands from DGGE profiles and intact polar lipid analyses were in accordance with the bacterial clone library results. Epsilonproteobacterial OTUs were affiliated with *Sulfuriconvexum*, *Arcobacter* and *Sulfurospirillum* spp. detected in other oil field habitats. The dominant organism revealed by the bacterial libraries (87% of all sequences) is a close relative of *Sulfuriconvexum kuijense* – an organism capable of oxidizing reduced sulfur compounds in crude oil. Geochemical analysis of organic extracts from bitumen at different reservoir depths down to the oil water transition zone of these oil sands indicated active biodegradation of dibenzothiophenes, and stable sulfur isotope ratios for elemental sulfur and sulfate in formation waters were indicative of anaerobic oxidation of sulfur compounds. Microbial desulfurization of crude oil may be an important metabolism for *Epsilonproteobacteria* indigenous to oil reservoirs with elevated sulfur content and may explain their prevalence in formation waters from highly biodegraded petroleum systems.

**Introduction**

The global inventory of petroleum reserves is dominated by heavy oil reservoirs that represent a legacy of anaerobic microbial communities that have degraded hydrocarbons *in situ* over geological timescales (Head et al., 2003; 2010; Larter et al., 2006). These biologically altered fossil fuels represent an enormous energy resource, yet their production is complicated by detrimental factors such as increased oil viscosity, acidity and sulfur content. Extreme cases of biodegradation in ‘unconventional’ super-heavy oil reservoirs have led to the development of novel extraction techniques involving mining or steam injection for producing this petroleum. A prominent example is the Athabasca oil sands (Fig. 1), where increased production...
in recent years underscores Canada having the second largest proven domestic crude oil reserves worldwide (c. 170 billion barrels; ERCB, 2009). The many challenges associated with producing heavy oil or bitumen (highly viscous oil) translate into added costs and increasing environmental impacts, and highlight the ongoing need for innovation in the oil sands sector. Biotechnologies based on understanding and manipulating subsurface reservoir microbial communities have been proposed as a promising route towards more sustainable production of energy from fossil fuels including oil sands (Grigoryan and Voor- douw, 2008; Jones et al., 2008; Youssef et al., 2009). These strategies are underpinned by a view of petroleum reservoirs as important habitats within the ‘deep biosphere’. Investigating reservoir habitats is often enabled by sampling produced water, which has been conducted for several oil fields including some in western Canada (Voordouw et al., 1996; Grabowski et al., 2005a). Similar studies from unconventional oil sands reservoirs have not been reported to date.

The occurrence of fermentative bacteria and methano- genetic archaea in several oil fields (Jeannot et al., 2005; Ollivier and Cayol, 2005) befits current models of anaerobic hydrocarbon degradation in subsurface petroleum reservoirs. Hydrocarbon geochemistry and gas isotope data have highlighted crude oil hydrocarbon degradation dominated by CO₂ reduction to methane as an important mechanism (Jones et al., 2008; Gray et al., 2009; 2010). Experimental microcosms inoculated with surface sediments show that anaerobic degradation of hydrocarbons and crude oil can be catalysed by consortia dominated by hydrogenotrophic (CO₂-reducing) methanogens (e.g. order Methanomicrobiales) and fermentative syntrophic Deltaproteobacteria (e.g. family Syntrophaceae) (Zengler et al., 1999; Jones et al., 2008). This hypothesis is consistent with the majority of the methanogenic archaea recovered so far from oil fields being hydrogenotrophic CO₂ reducers (Jeannot et al., 2005; Nazina et al., 2006; Gray et al., 2009). Interestingly, bacterial community analyses of reservoir samples have not revealed a clear prominence of Syntrophaceae to mirror the widespread occurrence of Methanomicrobiales. The absence, or low abundance of these putative syntrophs in oil reservoir communities may indicate that other bacterial groups might fill this niche in subsurface petroleum ecosystems. An alternative explanation could be the relatively few
investigations of indigenous microbial communities from actively biodegrading oil reservoirs where syntrophs are expected to be important in situ. Grabowski and colleagues (2005a,b) examined formation waters (i.e. co-produced with crude oil during primary recovery operations prior to secondary water flooding) from the Pelican Lake oil field – a 400 m deep conventional heavy oil reservoir in west Athabasca (Fig. 1). Fatty acid-degrading Syntrophus spp. could be enriched from this formation water after several transfers (Grabowski et al., 2005b), but were not detected in a bacterial 16S rRNA gene clone library constructed from DNA extracted from formation water (Grabowski et al., 2005a). This clone library consisted of 151 sequences from a single taxon affiliated with the genus Arcobacter within the Epsilonproteobacteria (Grabowski et al., 2005a).

Even though Epsilonproteobacteria have been detected in several oil fields using a variety of methods (e.g. Voordouw et al., 1996; Telang et al., 1997; Gevertz et al., 2000; Grabowski et al., 2005a; Hubert and Voor- douw, 2007; Sette et al., 2007; Gittel et al., 2009; Pham et al., 2009), the unexpected predominance of sequences from Epsilonproteobacteria in the Pelican Lake clone library has led to the suggestion that this may represent an artefact arising from selective amplification of epsilonproteobacterial 16S rRNA gene sequences by certain PCR primers (Grabowski et al., 2005a; Pham et al., 2009; Head et al., 2010). Such primer selectivity was shown to explain similar results obtained by Watanabe and colleagues (2000) who reported strong dominance of Epsi- lonproteobacteria in a similar environment (91% of clones derived from groundwater in an underground crude oil storage cavity), but later reported that using different PCR primers resulted in recovery of a greater diversity of sequences and a smaller proportion affiliated with Epsilonproteobacteria (no more than 16% of clones; Watanabe et al., 2002).

Here we present a microbiological characterization of formation waters from an unconventional shallow oil sands reservoir (0–80 m) from the northern Athabasca region in western Canada where the oil is severely degraded and immobile. Geochemical, microbiological, and lipid- and DNA-based molecular analyses of crude oil and formation water were performed, including PCR-based screening of 16S rRNA gene sequences from resident Archaea and Bacteria by employing different DNA extraction and PCR protocols to circumvent selectivity imposed by using only a single approach. The results reveal a strong dominance of Epsilonproteobacteria in the formation water samples from this oil sands reservoir. Coupling this observation with geochemical data showing that dibenzothiophenophanes are extensively biodegraded throughout the oil column and particularly at the oil water transition zone, we suggest that the predominance of Epsilonproteobacteria may be explained by the ability of these bacteria to use organic sulfur compounds in crude oil as an electron donor and energy source (Kodama and Watanabe, 2003).

### Results

#### Site description and formation water geochemistry

Samples were obtained from the Muskeg River mine, located approximately 75 km north of the city of Fort McMurray, Alberta (Fig. 1). In this oil sands reservoir, formation water occurs as a basal aquifer up to 20 m in thickness underlying a 50–80 m thick layer of oil-saturated sands. Formation waters are therefore not in contact with surface waters. Muskeg River oil sands are produced by clearing several metres of topsoil (overburden) to allow large-scale truck and shovel operation to excavate and transport oil sands to nearby facilities that separate the bitumen from the sand. This unconventional surface mining approach requires advanced dewatering of the oil sands to reduce formation pressure and prevent fractures and flooding as the open pit excavation proceeds deeper into the reservoir. Therefore formation waters are discharged at dewatering wells located 500–2500 m ahead of the advancing production area. Discharged formation waters should thus derive from biologically active oil water transition zones and associated aquifers in the subsurface (Head et al., 2003; Larter et al., 2003) and represent pristine reservoir samples unaffected by any prior water or chemical injections that sometimes characterize produced water samples from deeper conventional oil fields (Magot, 2005). Waters discharged from six wells around the mine excavation area had a combined flow rate of 250 m^3^ h^-1^, and three of these wells were sampled for geochemical and microbiological analyses (Fig. 1). The formation water geochemistry is summarized in Table 1, and indicates that the three samples are broadly similar. Similar microbial community compositions in the three water samples were confirmed by denaturing gradient gel electrophoresis (DGGE) (not shown) and intact polar lipid (IPL) analyses (Oldenburg et al., 2009). Further community analyses focused on formation water from wellhead #1.

#### Formation water microbial community composition

Two archaeal clone libraries (one prepared in Newcastle and the other in Calgary and Rockville) and three bacterial clone libraries (two prepared in Newcastle and one in Calgary and Rockville) were constructed from DNA that was extracted from the biomass that had been concentrated by filtering 12 l of formation water from wellhead #1. The diversity in all libraries was low with just seven
archaeal and five bacterial phylotypes being recovered from the five libraries (Table 2; operational taxonomic units (OTUs) were defined at 97% sequence identity). Clone library results indicate that regardless of DNA extraction procedure or PCR primers used, similar community compositions were obtained for Archaea and Bacteria in the oil sands formation water. The Newcastle archaeal library (91 clones) consisted of just a single phylotype affiliated with *Methanoregula* (Table 2). This phylotype was also dominant (80% of clones) in the larger community compositions were obtained for *Archaea* and *Bacteria* in the oil sands formation water. The Newcastle archaeal library (91 clones) consisted of just a single phylotype affiliated with *Methanoregula* (Table 2). This phylotype was also dominant (80% of clones) in the larger

Table 1. Geochemistry of oil sands formation water samples.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Wellhead #1</th>
<th>Wellhead #4</th>
<th>Wellhead #6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basal aquifer thickness (m)</td>
<td>18.6</td>
<td>17.6</td>
<td>23.0</td>
</tr>
<tr>
<td>Cumulative discharge at the time of sampling (m3 h⁻¹)</td>
<td>227 000</td>
<td>373 000</td>
<td>387 000</td>
</tr>
<tr>
<td>Discharge flow rate (flow continuity) (continuous)</td>
<td>10–20</td>
<td>20–30</td>
<td>60–70</td>
</tr>
<tr>
<td>Electrical conductivity (mS cm⁻¹)</td>
<td>2.36</td>
<td>5.22</td>
<td>4.20</td>
</tr>
<tr>
<td>pH</td>
<td>7.20</td>
<td>7.38</td>
<td>7.43</td>
</tr>
<tr>
<td>Alkalinity (mg L⁻¹ HCO₃⁻)</td>
<td>959.3</td>
<td>1512.0</td>
<td>1387.0</td>
</tr>
<tr>
<td>Intact polar lipids detected</td>
<td>Phosphatidylethanolamine; phosphatidylglycerol</td>
<td>Phosphatidylethanolamine; phosphatidylglycerol</td>
<td>Phosphatidylethanolamine; phosphatidylglycerol</td>
</tr>
<tr>
<td>δ²⁸S elemental S (%)</td>
<td>ND</td>
<td>26.7</td>
<td>26.7</td>
</tr>
<tr>
<td>S²⁻ SO₄⁻ (mg L⁻¹)</td>
<td>22.8</td>
<td>23.6</td>
<td>23.4</td>
</tr>
<tr>
<td>Na⁺</td>
<td>454.00</td>
<td>1276.00</td>
<td>1033.10</td>
</tr>
<tr>
<td>Cl⁻</td>
<td>351.22</td>
<td>1226.11</td>
<td>877.24</td>
</tr>
<tr>
<td>Br⁻</td>
<td>5.71</td>
<td>9.68</td>
<td>6.60</td>
</tr>
<tr>
<td>F⁻</td>
<td>3.23</td>
<td>3.60</td>
<td>3.72</td>
</tr>
<tr>
<td>Si⁶⁺</td>
<td>9.24</td>
<td>4.33</td>
<td>5.46</td>
</tr>
<tr>
<td>Ba²⁺</td>
<td>0.81</td>
<td>0.16</td>
<td>0.22</td>
</tr>
<tr>
<td>Sr²⁺</td>
<td>2.04</td>
<td>1.84</td>
<td>1.63</td>
</tr>
<tr>
<td>Ca²⁺</td>
<td>77.07</td>
<td>51.11</td>
<td>52.78</td>
</tr>
<tr>
<td>Mg²⁺</td>
<td>43.42</td>
<td>31.84</td>
<td>30.70</td>
</tr>
<tr>
<td>K⁺</td>
<td>18.10</td>
<td>22.10</td>
<td>19.40</td>
</tr>
</tbody>
</table>

Table 2. Archaeal and bacterial 16S rRNA gene clone library results.

<table>
<thead>
<tr>
<th>Library; target group (16S primers) (DNA extraction)</th>
<th>No. of clones sequenced in total</th>
<th>Lineages of phylotypes detected</th>
<th>% abundance in library (No. of clones)</th>
<th>Representative type sequence</th>
<th>Accession No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 Archaea (arch8f/arch1492r) (Calgary)</td>
<td>323</td>
<td>Methanoregula</td>
<td>88.9% (287)</td>
<td>TS1A275</td>
<td>JF789587</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Methanosipillum</td>
<td>9.3% (30)</td>
<td>TS1A121</td>
<td>JF789588</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Methanomicrobiales</td>
<td>0.6% (2)</td>
<td>TS1A142</td>
<td>JF789589</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Methanosarcina</td>
<td>0.3% (1)</td>
<td>TS1A083</td>
<td>JF789590</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Desulfovocculales</td>
<td>0.3% (1)</td>
<td>TS1A038</td>
<td>JF789591</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thermoplasmatelae 1</td>
<td>0.3% (1)</td>
<td>TS1A251</td>
<td>JF789592</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thermoplasmatelae 2</td>
<td>0.3% (1)</td>
<td>TS1A042</td>
<td>JF789593</td>
</tr>
<tr>
<td>A2 Archaea (arch46f/arch1017r) (Newcastle)</td>
<td>91</td>
<td>Methanoregula</td>
<td>100% (91)</td>
<td>(TS1A275)ᵣ</td>
<td>(JF789587)</td>
</tr>
<tr>
<td>B1 Bacteria (9f/1545r) (Calgary)</td>
<td>382</td>
<td>Sulfuricurvum</td>
<td>95.6% (369)</td>
<td>TS1B301</td>
<td>JF789594</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arcobacter</td>
<td>2.9% (11)</td>
<td>TS1B220</td>
<td>JF789595</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sulfurospirillum 1</td>
<td>0.3% (1)</td>
<td>TS1B252</td>
<td>JF789596</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pelibacter</td>
<td>0.3% (1)</td>
<td>TS1B322</td>
<td>JF789597</td>
</tr>
<tr>
<td>B2 Bacteria (8f/1542r) (Newcastle)</td>
<td>72</td>
<td>Sulfuricurvum</td>
<td>63.8% (46)</td>
<td>(TS1B301)</td>
<td>(JF789594)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arcobacter</td>
<td>31.9% (23)</td>
<td>(TS1B200)</td>
<td>(JF789595)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sulfurospirillum 2</td>
<td>4.2% (3)</td>
<td>NCL08_D6E05</td>
<td>JF789598</td>
</tr>
<tr>
<td>B3 Bacteria (ino-341f/1492r) (Newcastle)</td>
<td>96</td>
<td>Sulfuricurvum</td>
<td>63.5% (61)</td>
<td>(TS1B301)</td>
<td>(JF789594)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Arcobacter</td>
<td>29.2% (28)</td>
<td>(TS1B220)</td>
<td>(JF789595)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sulfurospirillum 2</td>
<td>6.3% (6)</td>
<td>(NCL08_D6E05)</td>
<td>(JF789598)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pelibacter</td>
<td>1.0% (1)</td>
<td>(TS1B322)</td>
<td>(JF789597)</td>
</tr>
</tbody>
</table>

a. Designations in parentheses indicate the type sequence is from one of the other libraries.

J. Craig Venter Institute (JCVI) library (n = 323 clones analysed), which contained additional phylotypes related to Methanospirillum (9% of clones), unclassified Methanomicrobiales (0.6%), Methanosarcina (0.3%), Desulfurococcales (0.3%) and Thermoplasmatales (0.3% for each of two distinct OTUs) (Table 2).

Common to all three bacterial libraries were two phylotypes related to Sulfitoferax and Arcobacter, with Sulfurospirillum always being the most abundant group (Table 2). Arcobacter was extracted using a bead beating method (30% Arcobacter clones in Newcastle libraries as opposed to 3% in the JCVI library; Table 2). Both of the Newcastle libraries revealed a distinct Arcobacter phylotype (4% and 6% of clones) that was not detected in the larger JCVI library, and the JCVI library contained a distinct Sulfurospirillum phylotype (a singleton) that was not detected in the Newcastle libraries. The only non-epsilonproteobacterial clones were singletons from the two larger libraries that comprised a fourth phylotype related to Pelobacter in the Delta proteobacteria.

Only two bacterial clone libraries were constructed initially, using primers 9f/1545r and 8f/1542r (Table S1), resulting in 99.7% and 100% of the cloned 16S rRNA gene sequences being affiliated with Epsilonproteobacteria (Table 2). This was unexpected, and similar to the 91% dominance of Epsilonproteobacteria in an underground storage cavity reported by Watanabe and colleagues (2000) using a similar primer pair (8f/1546r). Watanabe and colleagues (2002) subsequently constructed an additional library using an inosine-substituted 341f primer together with 1492r, which resulted in only 16% of clones affiliated with Epsilonproteobacteria (also related to Sulfuricurvum, Arcobacter and Sulfurospirillum). However, when the primers designed by Watanabe and colleagues (2002) were used to construct a third bacterial clone library from the oil sands formation water, again 99% of sequences were affiliated with Epsilonproteobacteria (Table 2).

The bacterial community composition determined by the three clone libraries was confirmed using a nested DGGE strategy whereby different universal bacterial primers (indicated above gel lanes; see also Table S1) was followed by a second round of amplification using DGGE primers to create amplicons 233 base pairs in length (including a 40-base-pair GC clamp) for DGGE, as described in the text. Arrows indicate bands from lanes 2–3 and 5–6 that were excised and sequenced. All sequences were closely related to the Sulfuricurvum OTU detected in the bacterial 16S rRNA gene clone libraries.

Archaea detected in oil sands formation waters

Putative hydrogenotrophic methanogens from the order Methanomicrobiales accounted for 410 out of 414 archaeal sequences retrieved from the oil sands formation water, comprising two dominant phylotypes related to Methanoregula and Methanospirillum spp. (Table 2; Fig. 3). Closely related Methanoregula spp. (> 98%; Fig. 3) have been detected in lake bottom sinkholes (e.g. Nold et al., 2010), a TCE-contaminated aquifer (Macbeth et al., 2004), a toluene-degrading enrichment (Ficker et al., 1999) and peatlands (e.g. Cadillo-Quiroz et al., 2008). One such peat bog was the source for the isolation of the type species Methanoregula boonei (Bräuer et al., 2006), an H2- and CO2-utilizing acidophilic methanogen that does not use other substrates (e.g. acetate, formate or trimethylamine; Bräuer et al., 2011) and is the closest cultured relative to the dominant oil sands phylotype (Fig. 3). The second most abundant phylotype has the hydrogenotrophic Methanospirillum hungatei as its closest cultured relative. Closely related Methanospirillum spp. have also...
been detected in sinkholes, aquifers and peatlands (see Fig. 3), and also include a Methanospirillum sp. enriched from the Pelican Lake oil reservoir c. 200 km to the west of these oil sands (Grabowski et al., 2005a; Figs 1 and 3). One out of the 414 sequences retrieved was affiliated with the Methanosarcinales (Table 2; Fig. 3) and is closely related to a Methanosarcina sp. detected in an archaeal 16S rRNA gene clone library from Pelican Lake formation water (without prior selection enrichment). Three more singletons from the

archaeal libraries were affiliated with non-methanogen sequences retrieved from cold seeps and sinkholes (Table 2; Fig. 3).

**Fig. 3.** Neighbour-joining phylogenetic tree indicating taxonomic affiliations of cloned archaeal 16S rRNA gene sequences from Athabasca oil sands formation waters. For each of the archaeal OTUs the most closely related sequences from GenBank and other related sequences of interest (e.g. from other petroleum reservoirs or methanogenic hydrocarbon-degrading systems) are indicated. Bootstrap values next to branching nodes are based on 100 resamplings.

Bacteria detected in oil sands formation waters

Bacterial sequences affiliated with Sulfuricurvum were most closely related to environmental clones derived from meromictic lake sediments (Nelson et al., 2007), chemoclines (unpublished GenBank entry Accession No. DQ82124).

Sulfurospirillum phylotypes discovered in this study share 94% 16S rRNA sequence identity with each other (Fig. 4). Sulfurospirillum phylotype 2 has among its closest relatives an organism enriched from the Pelican Lake formation water in medium selective for fermentative bacteria (Grabowski et al., 2005a). Other relatives include uncultured bacteria from a TCE-contaminated subsurface aquifer (Macbeth et al., 2004) and soil associated with a Himalayan glacier (unpublished GenBank entries, e.g. Accession No. EU809872; Fig. 4). The closest cultivated relatives of both Sulfurospirillum phylotypes are Sulfurospirillum multivorans and S. halo- respirans, which are metabolically versatile facultative chemolithotrophs (Luijten et al., 2003; Fig. 4). Sulfurospirillum phylotype 1 is most closely related (94%) to an uncultured lacustrine organism (unpublished GenBank entry Accession No. FJ612310; Fig. 4). Sulfurospirillum spp. have been detected in different oil field habitats, including the Kuji oil cavity, Pelican Lake and other fields in western Canada (Grabowski et al., 2005a; Hubert and Voordouw, 2007; Kodama et al., 2007; Fig. 4).

The Arcobacter spp. are found in various habitats (Gevertz et al., 2000; Campbell et al., 2006; Miller et al., 2007; Fedorovich et al., 2009; Webster et al., 2010), and those closely related to the oil sands phylotype come from a range of environments (Fig. 4), including the Kuji oil cavity (Watanabe et al., 2000). The most closely related cultured Arcobacter (strain R-28314; Fig. 4) was isolated from a denitrifying wastewater treatment plant (Heylen et al., 2000). An Arcobacter that made up 100% of the bacterial 16S rRNA gene clone library in Pelican Lake formation waters (Grabowski et al., 2005a) shares 92% 16S rRNA sequence identity with the Arcobacter detected in the oil sands formation waters studied here (Fig. 4).

Microbial community analysis of the formation water samples also included isolation and characterization of IPLs from biomass that was filtered from the formation water samples. Phosphatidylethanolamines (PE) and phosphatidylglycerols (PG) were detected in all three samples (Table 1; for further details see Oldenburg et al., 2009), indicating that formation waters retrieved from distant dewatering wells (kilometres apart; Fig. 1) may harbour similar bacterial communities. IPL structures can be used to make taxonomic inferences (Sturt et al., 2004), and while this approach lacks the phylogenetic resolution of 16S rRNA-based comparisons, bulk IPL analyses are not prone to issues such as primer specificity that can result in selective amplification of 16S rRNA genes from certain taxa. PE and PG are not highly diagnostic, as they are characteristic of many bacteria including Proteobacteria (Dowhan, 1997). In a recent study PE and PG were reported to be the most abundant IPLs in the anoxic zone of the Black Sea water column (Schubotz et al., 2009) where nucleic acid-based cloning and fluorescence in situ hybridization analyses both revealed dominance of Epsilonproteobacteria (Vetriani et al., 2003; Lin et al., 2006), and high rates of sulfide oxidation are coupled to nitrate reduction (Jørgensen et al., 1991; Wakeham et al., 2007). We did not detect IPLs characteristic of archaea in oil sands formation waters, suggesting that bacteria are more abundant than archaea in these samples.

**Sulfur biogeochemistry in oil sands reservoir formation waters**

During sampling of the reservoir formation waters, a distinct smell of hydrogen sulfide was noticed. The hydrogen sulfide concentration in the formation water samples was not measured during field sampling, but later chemical analysis detected up to 0.3 mM sulfate in the water samples (Table 1). Sulfur isotope analysis revealed mean δ34S values for elemental sulfur and sulfate of +26.7‰ and +23.3‰, respectively (Table 1), suggesting sulfate generation via microbial oxidation of elemental sulfur and possibly other reduced sulfur compounds. This 4‰ difference in δ34S between sulfur and sulfate is consistent with isotope fractionation effects observed in a pure culture of sulfide-oxidizing Epsilonproteobacteria (Sulfurimonas sp. strain CVO) from a western Canadian oil field, grown under nitrate-reducing conditions (Hubert et al., 2009).
various oil field Sulfurospirillum spp. (e.g., Grabowski et al., 2005a; Hubert & Voordouw, 2007)

- oil sands *Sulfurospirillum* 1 (0.2% of clones) JF789596
  - Sulfurospirillum caveoli isolated from Kuji oil storage cavity (Kodama et al., 2007) AB246781
  - Pelican Lake fermentative enrichment clones (Grabowski et al., 2005a) e.g., clone PL-26B3; AY570661
  - Sulfurospirillum multivorans (T) (Scholz-Muramatsu et al., 1995) X82931
  - Sulfurospirillum halorespians (Luijten et al., 2003) NR_028771
  - Kuji oil storage cavity clone 1014 (Watanabe et al., 2000) AB030587
  - glacier soil clones (unpublished) e.g., clone KS-31B; EU809872
  - glacier soil clone KF-7 (unpublished) EF445480
  - TCE-contaminated groundwater clone TAN6B (Macbeth et al., 2004) AY687251

- oil sands *Sulfurospirillum* 2 (1.6% of clones) JF789598
  - methane seep enrichment clone MnrB-B13 (Beal et al., 2009) FJ264592
  - Benzene-degrading enrichment clones (Herrmann et al., 2010) e.g., clone ZZ-54D12; EF613424
  - Arcobacter strain FWK08 isolated from a Canadian oil field (Gevertz et al., 2000) AF144693
  - Kuji oil storage cavity clones (Watanabe et al., 2000) e.g., clone 1006; AB030597
  - estuarine sediment clone 1CSR2-R1 (Webster et al., 2010) FN424330
  - Western Canadian oil field clone ARCO4 (Voordouw et al., 1996) U460505
  - Pelican Lake oil field formation water clones (Grabowski et al., 2005a) e.g., clone PL-8B1; AY570593
  - Alaska North Slope oil field clone D004025C09 (Pham et al., 2009) EU721619
  - Arcobacter strain ED-1 isolated from a microbial fuel cell (Fedorovich et al., 2009) FJ966634
  - Brazilian offshore oil field clone bo13C09 (Sette et al., 2007) AY962492
  - North Sea oil field clone DanBac25 (Gittel et al., 2009) FN356234
  - Black Sea sediment clone Ajb2 (Thamdrup et al., 2000) AJ271655
  - Arcobacter strain R-28314 isolated from a dentiflving wastewater (Heylen et al., 2006) AM084114
  - wastewater treatment plant biofilm clone i56 (Fernandez et al., 2008) AY692044
  - Alaskan North Slope oil field clone D004022A07 (Pham et al., 2009) EU721785

- oil sands *Arcobacter* (11% of clones) JF789595
  - vegetable harvest processing clone ATB-LH-6148 (unpublished) FJ535178
  - human liver carcinoma clone (unpublished) AF246706
  - Kuji oil storage cavity clone 1053 (Watanabe et al., 2000) AB030593
  - Sulfurimonas strain CVO isolated from a Canadian oil field (Gevertz et al., 2000) U460506
  - estuarine sediment clone 1CSR2-R8 (Webster et al., 2010) FN424332
  - gold mine water microbial mat clone ZSB-D4-6 (unpublished) GU025603
  - Sulfuricurvum kujense strain YK-1 isolated from Kuji oil storage cavity (Kodama & Watanabe, 2003) AB053951
  - Frassassi cave spring clone FC5A, EPS2, c30 (Porter & Engel, 2008) DQ295663
  - Frassassi cave water clone PC06110_B17 (Macalady et al., 2008) EU101169
  - Kuji oil storage cavity clones (Watanabe et al., 2000) e.g., clone 1005; AB030600
  - Sulfurospirillum strain CVO isolated from a Canadian oil field (Gevertz et al., 2000) U460506

- oil sands *Sulfuricurvum* (87% of clones) JF789594
  - meromictic lake water clone SLB117 (Nelson et al., 2007) DQ787677
  - meromictic lake chemocline clone eub26B (unpublished) GQ390209
  - aphotic cave sulfidic springs clone LKC3_127B.26 (Engel et al., 2004) AY510177
  - meromictic lake chemocline clone HH285 (Halm et al., 2009) FJ502274
  - clones from underground caves and springs (Porter & Engel, 2008) e.g., clone CC1_cL51; DQ295597
  - Sphingomonas strain LYP (Liu et al., 1999) AF126282
  - hexadecane-degrading methanogenic enrichment clone B1 (Zengler et al., 1999) AJ133794
  - Pelican Lake oil field methanogenic enrichment clone (Grabowski et al., 2005b) AY586399
  - North Sea oil field formation water clone TCB131x (Dahle et al., 2008) DQ647155
  - Dutch onshore oil field clone c95Bk26b (van der Kraan et al., 2010) FJ941547
  - Pelican Lake oil field formation water clone TCB131x (Dahle et al., 2008) DQ647155
  - heavy oil contaminated aquifer clone D12_07 (Winderl et al., 2008) EU266817
  - freshwater seep clone IS-83 (Bruun et al., 2010) GQ339169
  - oil sands *Pelobacter* (0.2% of clones) JF789597
  - pharmaceutical wastewater clone C11 (unpublished) EU234232
  - meromictic lake chemocline clone HH285 (Halm et al., 2009) FJ502274

Fig. 4. Neighbour joining phylogenetic tree indicating taxonomic affiliations of cloned bacterial 16S rRNA gene sequences from Athabasca oil sands formation waters. For each of the bacterial OTUs the most closely related sequences from GenBank and other related sequences of interest (e.g. from other petroleum reservoir habitats) are indicated. Bootstrap values next to branching nodes are based on 100 resamplings.
To assess the possibility that organic sulfur compounds provide electron donors for oil-associated Epsilonproteobacteria (as reported for S. kuijense strain YK-1; Kodama and Watanabe, 2003) dibenzothiophene concentrations were measured at different depths in drill core samples obtained from the oil sands layer overlying the formation waters. Figure 5 shows the dibenzothiophene concentration in oil decreasing from 215 ppm near the surface to 4 ppm near the oil–water contact at 80 m. This profile mirrors the general trends of oil biodegradation with depth, as shown by the concentrations of C_0 to C_6 alkyl naphthalenes, which decrease from 1011 ppm to 12 ppm in the same samples (alkyl aromatic hydrocarbon distributions and concentrations are indicative of the extent of biodegradation in heavy oils; Head et al., 2003; Larter et al., 2008). These depth profiles are consistent with microbial activity occurring over geological timescales at the oil water contact (c. 80 m depth in this system; further details provided in Fustic et al., 2011), and suggest that microbial desulfurization of crude oil compounds has occurred in these oil sands.

**Discussion**

An analysis of 19 clone libraries of bacterial 16S rRNA genes from different oil fields revealed that Epsilonproteobacteria are among the most frequently detected taxa in fossil fuel reservoir production fluids (Fig. 6). On the basis of these published reports, only Firmicutes and Gammaproteobacteria have been detected in oil fields more often than Epsilonproteobacteria. Grouping these 19 studies according to low-temperature (<50°C) and high-temperature (>50°C) reservoirs reveals that in low-temperature systems like the near-surface oil sands reservoir investigated here, Epsilonproteobacteria are the most frequently reported bacterial group (Fig. 6A). Furthermore, in low-temperature oil fields in western Canada, Epsilonproteobacteria are represented at high abundance in 16S rRNA gene clone libraries (Fig. 6B), as demonstrated here for the Athabasca oil sands (Table 2). By applying multiple bacterial PCR primers to DNA extracted using different techniques, our study confirms that Epsilonproteobacteria are indigenous to low-temperature petroleum reservoirs where they are abundant members of the microbial community; their massive dominance inferred from several 16S rRNA gene-based analyses is unlikely to be the result of primer selectivity as has been previously suggested (Grabowski et al., 2005a; Pham et al., 2009; Head et al., 2010). The oil sands at the Muskeg River mine are part of the same Athabasca heavy oil system as the Pelican Lake reservoir (Fig. 1). The recovery of 16S rRNA gene sequences from Epsilonproteobacteria (99–100% of clones) in pristine formation waters from both oil fields is in good agreement, and suggests that the dominance of Arcobacter spp. in the Pelican Lake formation waters (Grabowski et al., 2005a) reflects an important role or roles for Epsilonproteobacteria in heavy oil reservoirs.

The most abundant phylotype in all three of our bacterial libraries belonged to the genus Sulfuricurvum. Unlike the other epsilonproteobacterial genera detected (Arcobacter and Sulfurospirillum spp.; Table 2; Fig. 4), Sulfuricurvum spp. are not known to be capable of chemoorganotrophic metabolism on the basis of substrate tests with the only cultivated member of this genus, S. kuijense strain YK-1 (Kodama and Watanabe, 2003; 2004). Despite being autotrophic, strain YK-1 grows particularly well by oxidizing reduced, most probably organic, sulfur compounds in crude oil. This trait may explain Sulfuricurvum dominating an oil sands environment with a relatively high sulfur content (Strausz and Lown, 2003). The oil–water contact in the Muskeg River oil sands shows distinct gradients of dibenzothiophenes indicative of biodegradation in the aquifer associated with the oil sands (Fig. 5; Head et al., 2003). Biological sulfur oxidation in these oil sands is apparent from the...
isotopic composition of sulfate (mean +23.3‰) being enriched in $^{32}$S relative to elemental sulfur (mean +26.7‰) in the formation waters, even in the absence of nitrate (Table 1) – the electron acceptor that supports growth of *S. kujiense* on organic sulfur compounds in oil (Kodama and Watanabe, 2003). Oxidation of reduced sulfur compounds *in situ* may have been driven by microbial reduction of metal species (Aller and Rude, 1988; Beal *et al.*, 2009), and the presence of Fe$^{2+}$ and Mn$^{2+}$ in the formation waters (Table 1) is indicative of metal cycling in this system. Many *Arcobacter* and *Sulfurospirillum* spp. have been physiologically characterized and some are capable of iron or manganese reduction (Thamdrup *et al.*, 2000; Fedorovich *et al.*, 2009; Fry *et al.*, 2009); however, *S. kujiense* strain YK-1 has not been tested for this capability. Like some other *Epsilonproteobacteria*, strain YK-1 can grow under microaerophilic conditions (Kodama and Watanabe, 2003; 2004), and although these oil sands formation waters were noticeably sulfidic during sampling, rapid metabolism of low levels of oxygen entering this system via groundwater flow cannot be excluded.

*Sulfuricurvum* spp. have been detected in many groundwater environments (Campbell *et al.*, 2006; Porter and Engel, 2008; Fig. 4). The Kuji underground crude oil storage cavity (the source of strain YK-1; Kodama and Watanabe, 2004) is not a pristine oil reservoir like the oil sands studied here, but rather an engineered subsurface habitat where temporarily stored crude oil comes into contact with the local groundwater (Watanabe *et al.*, 2000; 2002). Detection in oil field bacterial clone libraries (references are given in B) from subsurface fossil fuel reservoirs were evaluated for the occurrence of dominant taxa. Seven of the habitats were low-temperature environments (reported as <50°C *in situ*) as indicated by the dashed line in (A), and 12 were high-temperature oil reservoirs (>50°C). (A) shows that *Epsilonproteobacteria* are the third most frequently occurring group overall and were detected in five out of seven low-temperature systems. The abundance of *Epsilonproteobacteria* in these oil field clone libraries is summarized in (B), which indicates that *Epsilonproteobacteria* are particularly abundant in low-temperature oil fields in western Canada. The hatched bar indicates the discrepancy in results reported from the Kuji oil storage cavity (Watanabe *et al.*, 2000; 2002) as described in the Introduction.

*Fig. 6.* Results from 19 published bacterial clone libraries (references are given in B) from subsurface fossil fuel reservoirs were evaluated for the occurrence of dominant taxa. Seven of the habitats were low-temperature environments (reported as <50°C *in situ*) as indicated by the dashed line in (A), and 12 were high-temperature oil reservoirs (>50°C). (A) shows that *Epsilonproteobacteria* are the third most frequently occurring group overall and were detected in five out of seven low-temperature systems. The abundance of *Epsilonproteobacteria* in these oil field clone libraries is summarized in (B), which indicates that *Epsilonproteobacteria* are particularly abundant in low-temperature oil fields in western Canada. The hatched bar indicates the discrepancy in results reported from the Kuji oil storage cavity (Watanabe *et al.*, 2000; 2002) as described in the Introduction.

![Graph showing the detection in oil field bacterial clone libraries](image)

Indeed, *Sulfuricurvum* spp. have not been reported in oil field microbial diversity studies (e.g. those indicated in Fig. 6) until now. The subterranean biogeography of several *Sulfuricurvum* spp. in groundwaters (Porter and Engel, 2008) suggests that a physical connection may exist between this Athabasca oil sands reservoir and the surrounding basal aquifer, which would influence the oil sands reservoir microbiota and biogeochemistry at oil water contact zones where biodegradation occurs (Head et al., 2003). The ways in which indigenous petroleum reservoir microbial communities establish are largely unknown, and groundwater infiltration – thought to influence microbial community composition in some coal deposits (Tseng et al., 1998; Schlegel et al., 2011) – may influence the resident microbiota in shallow western Canadian biodegraded oil sands.

Hydrogenotrophic methanogenesis is thought to be responsible for the formation of heavy oil (Jones et al., 2008), and depends on CO₂ and H₂ production via fermentation of crude oil components. It remains unclear which taxa catalyse syntrophic fermentations in biodegraded subsurface petroleum reservoir environments (Pham et al., 2009), and the predominance of Epsilonproteobacteria is intriguing in this context. Certain studies suggest Epsilonproteobacteria might be involved in syntrophic anaerobic communities that degrade hydrocarbons. For example, an epsilonproteobacterium was recently shown to assimilate 13C during syntrophic degradation of 13C-labelled benzene in sulfate-reducing enrichment cultures (Herrmann et al., 2010) and appears to contribute to benzene fermentation, e.g. into H₂, acetate and CO₂. Indeed in the study of Herrmann and colleagues (2010), the highest 13C enrichment was observed for the Epsilonproteobacteria. Sulfurospirillum spp. can ferment a wide variety of organic compounds (Luijten et al., 2003) and were among the fermentative heterotrophs isolated from a subsurface coal deposit (Fry et al., 2009) and the Pelican Lake oil reservoir (Grabowski et al., 2005a; Fig. 4). Acetate is metabolized by Arcobacter spp. in various anoxic settings (Thamdrup et al., 2000; Fedorovich et al., 2009; Webler et al., 2010). Syntrophic oxidation of acetate to H₂ and CO₂ is important in some methanogenic crude oil degrading systems (Jones et al., 2008), and is catalysed by a few cultured isolates (Hattori, 2008; Westerholm et al., 2010) that can also catalyse the reverse reaction (i.e. acetogenesis from H₂ and CO₂; Lee and Zinder, 1988). Several Epsilonproteobacteria can use H₂ as an electron donor (e.g. Gevertz et al., 2000; Kodama et al., 2007), including the acetogenic Arcobacter dominating the bacterial community in the Pelican Lake reservoir (Grabowski et al., 2005a). This is consistent with Arcobacter sequences dominating the metagenome of formation water from a coalbed methane reservoir (US Patent 2010/0047793 A1, 2010), a habitat where syntrophic methanogenic hydrocarbon degradation also occurs (Strápač et al., 2008).

Intact polar lipid analyses indicated that bacteria predominated over archaea in the formation waters (Oldenburg et al., 2009). Taken together, lipid and DNA-based results therefore suggest a Sulfiticurvum sp. is the most abundant organism in this heavily biodegraded oil reservoir, with Arcobacter, Sulfurospirillum and hydrogenotrophic Methanomicrobicaceae spp. making up smaller fractions of the microbial community (Table 2). The latter groups may be involved in methanogenic crude oil degradation that has generated these and other heavy oils, and currently these populations may have given way to a dibenzothiophene-degrading Sulfiticurvum population well suited to this sulfur-enriched oil sands reservoir (around 5% sulfur by weight; Marcano, 2011). Our results provide clear evidence that 16S rRNA gene-based detection of Epsilonproteobacteria in oil field formation waters is not artefactual. On the contrary, Epsilonproteobacteria are predominant members of low-temperature biodegraded petroleum reservoirs, and their ecophysiology in such environments can be explained by examining different aspects of crude oil sulfur and hydrocarbon biogeochemistry.

Experimental procedures

Sample collection

Athabasca oil sands formation waters were sampled at the Muskeg River oil sands mine north of Fort McMurray, Alberta, Canada (Fig. 1). Mining operations at this site require reservoir formation water discharge in advance of open pit oil sands removal and water from three out of the six dewatering wells surrounding the active excavation area were sampled (Fig. 1). Wellhead valves were opened and flushed for several minutes prior to collecting 12 or 24 l using a sterilized metal funnel to completely fill sterile 4 l brown glass bottles that had previously been flushed with nitrogen gas. Water samples were transported to the laboratory in a cool box where they were subsequently stored at 4°C. Oil sands samples were obtained from areas of the reservoir that had been undergoing commercial excavation. Large volumes (several m³) of sediment were excavated from the intact formation by mine equipment (large backhoe). This enabled aseptic subsampling from the middle part of the excavated oil sands. Subsamples were stored anaerobically in N₂-flushed containers kept at room temperature. Oil sands were also obtained from vertical drill cores spanning the oil leg and reaching depths close to the oil water transition zone.

Geochemical analysis

Formation water samples from dewatering wells numbered 1, 4 and 6 were subject to a suite of geochemical analyses, performed by the Analytical Geochemistry Group at the University of Calgary. To determine isotopic compositions of sulfate and elemental sulfur in the formation waters, BaSO₄ was precipitated following addition of 0.5 M BaCl₂, and elemental sulfur was extracted following filtration and Bligh and Dyer extraction (Bligh and Dyer, 1959). Sulfur isotope ratios were determined by continuous flow-elemental analysis-isotope ratio mass spectrometry (CF-EA-IRMS) with a Finnigan Mat Delta+XL spectrometer interfaced with a Carlo Erba NA 1500 elemental analyser and are expressed relative to the international standard Canyon Diablo Troilite. For IPL analysis, biomass was concentrated by vacuum filtration of the 12 or 24 l samples through 0.2 μm pore size filters (Millipore) that were subsequently freeze dried to allow biomass removal by scraping it from the filters with a sterilized metal spatula. The biomass was then subjected to a modified Bligh and Dyer extraction (Bligh and Dyer, 1959) to recover polar fractions which were analysed by liquid chromatography-mass spectrometry to determine IPL composition, as described previously (Oldenburg et al., 2009).
Bitumen samples from the oil sands drill core were subjected to petroleum geochemistry analyses. ‘Aromatic hydrocarbon’ fractions including dibenzothiophenones were isolated from de-asphalted bitumen extracts following the procedure of Bennett and Larer (2000). Alkyl naphthalenes were quantified relative to a D8-naphthalene standard (response factors of unity were employed therefore data can be considered semi-quantitative). Mass spectral characterization of compounds in the aromatic hydrocarbon fractions was carried out using splitless injection GC MS with an Agilent 6890 chromatograph interfaced to a 5973 quadrupole mass-selective detector. Saturated and aromatic hydrocarbons were analysed with a DB-5 coated fused silica column (30 m length, 0.32 mm id, 0.25 μm film thickness; J&W Scientific) according to the following temperature programme: 40°C (held for 2 min) to 300°C (held for 20 min) increasing at 4°C min⁻¹.

DNA extraction and PCR amplification of 16S rRNA genes

Most of the biomass isolated by filtration was dedicated to IPL analysis as described above. DNA was extracted from the remaining biomass residue on the same filters, which allowed for 16S rRNA gene analyses. Initial screening was performed by DGGE to confirm that the microbial communities in the three water samples were similar. Subsequent community analyses were performed on water from wellhead #1 according to two complementary approaches:

(i) At the University of Calgary, filters were submerged in a resuspension buffer (0.15 M NaCl; 0.1 M EDTA) and incubated on a rotary shaker overnight. DNA was extracted from resuspended biomass using the method of Marmur (Marmur, 1961) with modifications (Vooroud et al., 1990) including successive rounds of freezing and thawing as well as lysozyme, RNase and proteinase K treatments. DNA concentrations were measured by spotting dilutions on a square of 1% agarose gel containing ethidium bromide (c. 10 μg ml⁻¹). The concentration of DNA was estimated under UV illumination by comparison with fluorescence of known amounts of phage lambda DNA. The presence of bacterial DNA was confirmed in all formation water samples by positive PCR amplification from all DNA extracts using primers 8f and 1406r (see Table S1). Extracted DNA from wellhead #1 formation water was sent to the JCVI in Rockville for amplification of 16S rRNA from Archaea using primers arch46f and arch1017r. 16S rRNA genes from Bacteria were amplified using two different primer pairs, either 8f and 1542r or insine-341f and 1492r (Table 2). All primer sequences and a summary of downstream procedures used to analyse the amplified 16S rRNA genes are listed in Table S1.

(ii) Parallel analysis was conducted at Newcastle University where DNA was extracted directly from the filters for each wellhead formation water sample using the FastDNA Spin Kit for soil (Q-BIOgene, UK). Filters containing biomass residue were cut into small sections and added to 2 ml tubes and DNA was extracted following the manufacturer’s guidelines, which include an initial bead-beating step using a RiboLyzer (Hybaid). 16S rRNA genes from Archaea were amplified by PCR using primers arch46f and arch1017r. 16S rRNA genes from Bacteria were amplified using two different primer pairs, either 8f and 1542r or insine-341f and 1492r (Table 2). All primer sequences and a summary of downstream procedures used to analyse the amplified 16S rRNA genes are listed in Table S1.

Denaturing gradient gel electrophoresis (DGGE)

Shorter 16S rRNA gene fragments were amplified for DGGE. For bacterial 16S rRNA genes a PCR product generated using primers 341f-GC and 534r was used directly for DGGE. Bacterial DGGE was also performed following nested PCR amplifications where PCR products generated with insine-341f and 1492r or 8f and 1542r were used as templates for a second PCR reaction with 341f-GC and 534r. A nested approach was also used to analyse archaeal 16S rRNA gene fragments by DGGE. PCR products generated with primers arch46f and arch1017r were used as template for a second round of PCR using primers arch341f-GC and Uni522r. All primer sequences are listed in Table S1. DGGE was conducted in 10% acrylamide gels with 30–60% denaturant, as described previously (Rowan et al., 2003). Acrylamide gels were run for 4 h at 200 volts using the D gene system (Bio-Rad, Hemel Hempstead, UK) and were subsequently stained for 30 min in SYBR green 1/1000 in 1× TAE. Stained gels were viewed under ultraviolet light and gel images were recorded using a Bio-Rad Fluor-50 Multimaguer (Bio-Rad, UK). Bands of interest were excised for sequencing.

Cloning, sequencing and phylogenetic analyses

Bacterial and archaeal 16S rRNA gene fragments were cloned using TOPO TA cloning kits (Invitrogen) according to the manufacturer’s instructions. Approximately 384 clones were analysed from the JCVI libraries, and approximately 96 clones were analysed from each of the Newcastle libraries. Nearly complete full-length 16S rRNA gene sequences (Escherichia coli positions 9–1545) were obtained from the JCVI libraries by sequencing using M13 cloning vector (Pruesse et al., 2007) and alignments were manually corrected using BioEdit (Hall, 1999). Neighbour joining phylogenetic trees were constructed using MEGA4 (Tamura et al., 2007). Sequence alignments for phylogenetic tree reconstruction included the top three hits obtained in BLAST and SeqMatch searches (multiple environmental sequences from the same study or habitat were considered as one hit).
to each of the oil sands phylotypes, as well as other sequences of interest. Representative sequences from the formation water clone libraries have been deposited in the GenBank database under Accession No. JF789587 to JF789598.

Meta-analysis of published oil field 16S rRNA bacterial clone library results

A literature survey revealed 19 published clone libraries of bacterial 16S rRNA genes from oil reservoir production fluids. Clone library results were collated to reveal the proportion of clones affiliated with major bacterial divisions. This analysis considered over 3600 sequences of cloned 16S rRNA genes, which we classified at the level of major phylogenetic groups (phylum/subphylum level). From these data, distributions based on the frequency of occurrence and average percentage representation of major taxa in clone libraries could be plotted (see Fig. 6). Only clone libraries that used broad specificity (‘universal’) primers for PCR amplification of 16S rRNA genes were included in the survey. Published DGGE analyses on oil field production fluid samples were not included in this analysis. Results were considered in the context of low- and high-temperature petroleum reservoirs using a cut-off of 50°C (in situ reservoir temperature as indicated in the publications). We recently used a similar approach to conduct a broad assessment of microbial community structure in petroleum-impacted environments including soils and surface sediments (Gray et al., 2010).

Acknowledgements

We thank Albian Sands Energy for permission to obtain samples and Yonnus Idris for field assistance. Financial support for this work was provided from the Bacchus II consortium (Agip ENI, BP/Amoco, Chevron Texaco, Conoco Phillips, StatoilHydro, Petrobras, Saudi Aramco, Shell, Total and Woodside), the Alberta Science and Research Authority, the Canadian Foundation for Innovation, the Natural Sciences and Engineering Research Council of Canada, and a European FP7 Marie Curie International Incoming Fellowship to C.H.

References


**Supporting information**

Additional Supporting Information may be found in the online version of this article:

**Table S1.** PCR primers used for microbial community analyses.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.