Lawrence Berkeley National Laboratory

Lawrence Berkeley National Laboratory

Title

Biocorrosive Thermophilic Microbial Communities in Alaskan North Slope Oil Facilities

Permalink

https://escholarship.org/uc/item/0g5238bh

Author

Duncan, Kathleen E.

Publication Date

2009-09-16

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

LBNL-XXXXX
Preprint

ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

Title | Biocorrosive Thermophilic Microbial Communities in Alaskan North Slope Oil Facilities

Author(s) | Kathleen E. Duncan, Lisa M. Gieg, Victoria A. Parisi, Ralph S. Tanner, Joseph M. Suflita, Susannah Green Tringe, Jim Bristow

Division | Genomics

Journal Name | Environmental Science & Technology

One-sentence summary: Anaerobic degradation of low molecular weight hydrocarbons support thriving thermophilic archaeal and bacterial communities that produce a wide variety of metabolites that can stimulate biocorrosion in Alaskan North Slope oil production facilities.

Title: Biocorrosive Thermophilic Microbial Communities in Alaskan North Slope Oil Facilities

Authors: Kathleen E. Duncan[1], Lisa M. Gieg[1,3], Victoria A. Parisi[1], Ralph S. Tanner[1], Joseph M. Suflita[1], Susannah Green Tringe[2], and Jim Bristow[2].

1. Dept. of Botany and Microbiology and the Institute for Energy and the Environment, University of Oklahoma, Norman, OK, USA. 2. DOE Joint Genome Institute, Walnut Creek, CA, USA. 3. current affiliation: Dept. Biological Sciences, University of Calgary, Calgary, Alberta, Canada.

Abstract

Corrosion of metallic oilfield pipelines by microorganisms is a costly but poorly understood phenomenon, with standard treatment methods targeting mesophilic sulfate-reducing bacteria. In assessing biocorrosion potential at an Alaskan North Slope oil field, we identified thermophilic hydrogen-using methanogens, syntrophic bacteria, peptide- and amino acid-fermenting bacteria, iron reducers, sulfur/thiosulfate-reducing bacteria and sulfate-reducing archaea. These microbes can stimulate metal corrosion through production of organic acids, CO_2, sulfur species, and via hydrogen oxidation and iron reduction, implicating many more types of organisms than are currently targeted. Micromolar quantities of putative anaerobic metabolites of C_1-C_4 *n*-alkanes in pipeline fluids were detected, implying that these low molecular weight hydrocarbons, routinely injected into reservoirs for oil recovery purposes, are biodegraded and provide biocorrosive microbial communities with an important source of nutrients.

Introduction

The U.S. possesses a network of over 2.3 million miles of pipelines that transmit about 75% of the nation's crude oil and 60% of refined products (http://www.corrosioncost.com/infrastructure/gasliquid/index.htm). Despite this importance, pipelines are not regularly considered in assessments of societal infrastructure needs (http://www.asce.org/files/pdf/reportcard/2005_Report_Card-Full_Report.pdf), but there is little doubt that these facilities are vulnerable and can deteriorate over time. Through-wall breaches due to corrosion are expensive problems in the oil industry that can result in explosions, product interruptions, hazardous chemical releases and environmental damage. Such was the case in the August 2006 Prudhoe Bay release on Alaska's North Slope (ANS) (http://www.usatoday.com/news/nation/2006-08-06-alaskan-oil-field_x.htm). The metabolic activities of microorganisms were implicated in this and other incidents of pipeline failure. In fact, it has long been known that microbes contribute to corrosion by multiple mechanisms *(1-3)*, yet biocorrosion is not a well-understood process. There is no consensus on the identity of specific microorganisms responsible for corrosion or how they function to catalyze such incidents, resulting in poorly targeted efforts to monitor and combat biocorrosion.

Following the pipeline breach at Prudhoe Bay, we obtained samples from an ANS field to assess the potential for biocorrosion via metabolic indicators and microbial community analysis. The geology and geochemistry of ANS fields have previously been described *(4)* and subsurface conditions are well within the range for microbial communities to thrive *(5)*. The facility produces oil, gas and water from multiple anaerobic and hot (average temperature 68°C) reservoirs and is typical of ANS oilfields that collectively have produced up to 16% of the U.S. domestic oil requirements for over 30 years (http://tonto.eia.doe.gov/oog/special/eia_sr_alaska.html). Fluids and gases from multiple production wells are collected in a central facility, from which oil is channeled to the Trans-Alaska Pipeline System. At the facility, low molecular weight hydrocarbons (mostly methane, with lesser amounts of C_2-C_4 *n*-alkanes) and water are reinjected into the oil-bearing formations to maintain pressure and facilitate oil recovery. Most pipelines are above ground and thermally insulated, so conditions inside the pipelines and processing facilities are anaerobic and hot. As required, seawater is treated with biocide and added to maintain formation pressures or during oil processing, thus introducing seawater chemistry, a lower temperature, and potentially marine microorganisms into oil reservoirs. To assess biocorrosion potential in the ANS field, we obtained fluid samples at well heads from production wells (producing oil, gas and water), from a water reinjection well following oil processing activities in a central facility (CF), 2 locations within the CF (a 1st stage separator and a coalescer), from a pipeline carrying fluids and gas to the CF, treated seawater, and fluids and solids scraped from the inner surface of the pipeline carrying treated seawater. Samples were prepared accordingly for metabolite evaluation or for molecular community analysis *(6).*

Bacterial Community Profiling

Despite oil production from several major reservoirs with different geological histories, the facility-wide bacterial community profiles at the ANS field showed striking similarities for three of the high temperature sites. Bacterial communities from production well 2L, the 1st stage separator (PS) and the coalescer (CO) exhibited a high degree of class-level similarity (Fig. 1; table S1A), and greater levels of genetic diversity and species richness (table S2) than did the archaeal or the other two bacterial libraries. Ten "core" taxa (defined as OTUs with 97% nucleotide sequence similarity) were found at all three sites and represented over 87% of the bacterial 16S rRNA gene sequences (Fig. 2; table S3). Fig. 2 also illustrates the taxa and number of sequences shared between any two of the sites as well as those unique to each site. The most abundant of the core sequences (2L: 83%, PS: 57%, and CO: 31%) has 97-99% identity to that of *Thermovirga lienii (7)*. *T. lienii* is a thermophilic anaerobe isolated from a North Sea oil well and described as a member of the Firmicute family Syntrophomonadaceae.

However, it has also been designated a member of the candidate division Synergistes *(8)* and sequences similar to those of *T. lienii* will be designated as "Synergistes" here. The type strain of *T. lienii* has an optimum growth temperature of 58°C and ferments certain amino acids, proteinaceous substrates and organic acids, producing ethanol, acetate, propionate, isovalerate/2-methylbutyrate, H_2, and CO_2 *(7)*. It can also reduce cystine and elemental sulfur to H_2S. Synergistes-associated sequences were abundant in an extensively biodegraded mesophilic ANS oil reservoir *(9)*. The majority were more similar to uncultured *Thermovirga* clones than to the type strain as expected from the temperature range favored by *T. lienii.*

The most abundant delta proteobacterial 16S rRNA gene sequences (2L: 4%, PS: 7%, CO: 18%) were similar to that of *Desulfomicrobium thermophilum (10)*, a sulfate-reducing bacterium isolated from a hot spring. Sulfate-reducing bacteria (SRB) have been routinely monitored by the oil industry because of their ability to produce H_2S from sulfate. However, many other core sequences were similar to those of organisms that produce sulfide through the reduction of elemental sulfur, thiosulfate, sulfate, or other sulfur oxyanions (e.g. *Thermosipho africanus/ T. geolei, Pelobacter carbinolicus, Desulfacinum subterraneum,* and *Thermodesulfobacterium commune*). Clone libraries based on *dsrAB* genes that code for an essential enzyme for sulfate reduction (dissimilatory (bi)sulfite reductase) were heavily dominated by sequences similar to those of the archaeal sulfate-reducer *Archaeoglobus fulgidus* (2P: 60%, PS and CO: >99%). These findings suggest that bacterial sulfate reduction makes only a minor contribution to sulfide production at the facility.

Core sequences similar to *Thermoanaerobacter pseudethanolicus* (and other

Thermoanaerobacter species), *Thermacetogenium phaeum*, or *P. carbinolicus* indicate the possible importance of iron-reduction and/or syntrophic metabolic interactions *(11)*. Anaerobic iron-reducing thermophiles in deep subsurface petroleum reservoirs have been previously demonstrated *(12)*. However, thermophilic strains of *Pelobacter* are not yet known, so its detection as a core taxon awaits further explanation.

In comparison to the 2L and CF samples, the production well 2P bacterial community had much lower diversity and species richness (Fig 1; table S2). The dominant (90%) bacterial 16S rRNA gene sequence is similar (97-99%) to that of moderately thermophilic, organic acid-utilizing, nitrate-reducing *Petrobacter* species *(13)*. The second-ranked (4.7%) sequence type is similar to *T. lienii* and sequences similar to the core taxa *T. pseudethanolicus* and Thermotogales were present in low abundance (table S1A). Which specific factors are responsible for the differences are unknown; well head temperatures were similar for 2P and 2L (avg. 48°C and 49°C respectively) with little variation for 9 months prior to sampling, however the wells draw from different formations. Collectively, all well and CF samples strongly resemble anaerobic thermophilic oil reservoir and well communities.

Archaeal Community Profiling

Sequences similar to those of hyperthermophilic Archaea, notably sulfate-reducing *Archaeoglobus* species, methane-producing *Methanothermobacter thermautotrophicus*, and H_2S-producing *Thermococcus* were abundant in the PS, CO, and 2P samples. More than 90% of the archaeal 16S rRNA gene sequences fell into the corresponding three families of Euryarchaeota (Fig. 3, tables S1B and S4). Crenarchaeota were only detected in the PS sample, which also exhibited the highest diversity and species richness of the three archaeal libraries (table S2). All three Euryarchaeota groups have frequently been detected in hot oil reservoirs and production fluids *(5)*. *Archaeoglobus* and *Thermococcus* enriched from a North Sea oil field grew at high temperatures on crude oil as the sole source of carbon and nutrients *(14)*. The same study also found *Archaeoglobus*-like cells in hyperthermophilic cultures enriched from ANS reservoirs.

Sequences of methanogens (approximately ¼ of the archaeal 16S rRNA gene sequences) were less abundant than those of fermentative and sulfate-reducing archaea. Most methanogenic sequences were related to those of hydrogen-utilizing *Methanothermobacter* species. Hydrogen-utilizing methanogens have been commonly found in hot oil reservoirs *(15)*. Approximately 12% of methanogenic sequences were 99.8% similar to that of "*Methermicoccus shengliensis"* (DQ787474, Methanosaetaceae). "*M. shengliensis"* strain ZC-1 *(16)* was isolated from oil-production water and has optimal growth at 65°C. Strain ZC-1 is not an acetoclastic methanogen, unlike other members of the Methanosaetaceae. In contrast to our results, acetoclastic methanogens were by far the most abundant archaea in a heavily biodegraded mesophilic North Slope oil reservoir *(9)*.

Targeted Cultivation and Seawater Pig Envelope Community Profiling

In agreement with the molecular analysis, *M. thermautotrophicus* was isolated as the numerically dominant (2.3/mL) hydrogen-using prokaryote from the 1st stage separator. A thermophilic *Anaerobaculum* sp. was the numerically dominant heterotroph cultured from the same sample *(17)*. Members of the genus *Anaerobaculum* ferment organic acids and peptides but also reduce thiosulfate, sulfur, and cysteine to H_2S *(18)* and are thus directly implicated in biocorrosion. However, all populations of culturable bacteria screened (SRB, anaerobic/facultative heterotrophs, hydrogen-users) were found in low numbers (2-4 cells/mL), implying that these organisms would be missed in most routine screening procedures.

The seawater pig envelope sample (SW) community profile was quite different from that of the archaea-rich production wells and the CF, primarily consisting of sequences similar to those of mesophilic and psychrophilic marine bacteria. We were unsuccessful in obtaining a small subunit ribosomal archaeal RNA gene library with archaeal primers although a bacterial 16S rRNA gene sequence library was successfully obtained from the same DNA sample. Populations of culturable bacteria were 10^3 to 10^6 /mL, with the numerically dominant organism and most abundant DNA sequence from the seawater 16S rRNA gene library most similar to γ-proteobacteria, *Pseudomonas stutzeri* and/or related species (Fig. 1, table S1). The numerically dominant culturable hydrogen-user was *Acetobacterium*. An *Acetobacterium* species has previously been isolated from marine environments *(19)*. Sulfate-reducing bacteria were estimated at 2.4×10^6/mL. In accord with the 16S rRNA library results, no *dsrAB* sequences similar to those of *Archaeoglobus* were obtained from the SW sample. 16S rRNA gene sequences similar to those of the deep-sea genera *Sulfurimonas* and *Arcobacter* (epsilon proteobacteria, table S1) were abundant in the SW sample but were not found in the production wells or CF samples. Only two sequences from the other four bacterial libraries, one from the PS and one from well 2L (both *Pseudomonas*) were as much as 97% similar to any of the sequences from the SW sample. Thus, mesophilic and psychrophilic marine microorganisms originating from seawater seem unlikely to be responsible for biocorrosion problems at high temperature sites. However, seawater can contribute increased levels of sulfate, manganese, or organic matter that could spur increased corrosive microbial activity. For example, it was noted that H_2S was not detected in one ANS oilfield until after seawater flooding was initiated *(20)*.

Metabolic Profiling

It is well established that biocorrosive organisms form complex surface assemblages where cells are imbedded in a matrix of biologically-produced extracellular polymeric substance (EPS) that forms a protective microenvironment *(1)*. However, the carbon source(s) supporting the formation of such surface-associated communities remain enigmatic. Clearly, the largest potentially available source of carbon to support microbial activity is the oil itself. Since hydrocarbons are known to be suitable substrates for anaerobes *(21)*, we suspected that the more water-soluble oil components, like benzene, toluene ethylbenzene and xylene isomers (BTEX) might be preferentially metabolized to support the diverse microbial communities detected at the facility. This prospect was explored by assaying for the signature metabolites associated with anaerobic oil biodegradation *(22-23)*. The identification of these intermediates implicates the parent hydrocarbons being metabolized.

Contrary to expectations, there was no evidence for the biodegradation of the most water-soluble BTEX components. However, putative low molecular weight alkylsuccinate metabolites associated with anaerobic *n*-alkane biodegradation were found facility-wide. Six of the eight central facility and production well samples collected contained 0.8-2.2 μM concentrations of low molecular weight (C_1-C_4) alkylsuccinates (Table 1). No signature hydrocarbon metabolites were found in the seawater samples. The identification of methyl-, ethyl-, propyl- and butylsuccinate suggested that the hydrocarbons routinely reinjected during normal oil recovery operations were being biologically oxidized by a fumarate addition reaction in a manner analogous to higher molecular weight *n*-alkanes (Table 1) *(22)*. The recycling of these gases to help maintain formation pressures occurred throughout the decades-long production history of the formation and we suspect that the requisite organisms were enriched over this long time period.

If the analogy to higher molecular weight *n*-alkane anaerobic metabolism is accurate, we predict the formation of a series of downstream branched and straight-chain fatty acid metabolites formed as a result of the presumed carbon skeleton rearrangement and subsequent decarboxylation of the alkylsuccinate intermediates *(24)*. Indeed, the expected downstream metabolites were also found in the same samples from the central facility, but none were found in the seawater samples (Table 1).

The detection of methylsuccinate in conjunction with the other low molecular weight alkylsuccinates is particularly evocative. This finding suggests that fumarate addition may represent an alternative to previously described mechanisms such as reverse methanogenesis for the anaerobic oxidation of methane *(25)*. Consistent with an alternate hypothesis, the methanogen and ∂ proteobacterial sequences were dissimilar to the genera described in other environments undergoing anaerobic methane oxidation *(26)*.

Implications for Biocorrosion

Our results suggest that temperature and hydrocarbon utilization are primary factors governing microflora species composition at the ANS facility. Contrary to the emphasis placed on mesophilic bacterial SRB by standard biocorrosion monitoring procedures, many of the organisms detected using molecular techniques and targeted isolation are thermophilic bacteria capable of reducing various sulfur oxyanions or hyperthermophilic sulfate-reducing archaea that produce H_2S. Methanogenic, fermentative, H_2-producing and H_2-utilizing physiologies were also common, unlikely to be detected using standard techniques, and could likewise stimulate corrosion. The similarity of core taxa in these samples and those from other thermophilic oil reservoirs and wells suggests that hydrocarbon-degrading, potentially corrosive microbes found in oil reservoirs will readily inoculate and proliferate in oil production facilities maintained at compatible temperatures. Such similarities also imply that pipeline integrity management programs might be able to differentially target a relatively few core taxa.

The detection of putative low molecular weight alkane metabolites throughout the hot oilfield facilities suggests that anaerobic hydrocarbon biodegradation is inherent and likely involved in supporting biocorrosive biofilms. Indeed the formation of relatively high concentrations of alkylsuccinates (Table 1; µM vs nM concentrations more typically found in fuel-contaminated aquifers, *22*) allows us to postulate that such acidic intermediates can directly contribute to biocorrosion processes. The subsequent metabolism of the compounds would eventually form acetate and CO_2, microbial products known to exacerbate corrosion of pipeline surfaces *(3, 27)*. Thus, these findings support the hypothesis that anaerobic hydrocarbon biodegradation processes in the oilfield environment can be an important factor in microbial influenced corrosion.

References

1. I. B. Beech, J. Sunner, *Curr. Opin. Biotechnol.* **15,** 181-186 (2004).

2. W. A. Hamilton, *Biofouling.* **19**, 65-76 (2003).

3. J. M. Sulfita, T. J. Phelps, B. Little, *Corrosion* **64**, 854-859 (2008).

4. W. D. Masterson, L. I. P. Dzou, A. G. Holba, A. L. Fincannon, L. Ellis, *Organic Geochem*. **32**, 411-441 (2001).

5. M. Magot, in *Petroleum Microbiology*, B. Ollivier, M. Magot, Eds. (ASM Press, Washington, D.C., 2005), pp. 21-34.

6. Materials and methods are available as supporting material on Science Online

7. H. Dahle, N. K. Birkeland, *Int J Syst Evol Microbiol*. **56**,1539-1545 (2006).

8. S. R. Vartoukian, R. M. Palmer, W. G. Wade, *Anaerobe* **13**, 99-106 (2007).

9. V. D. Pham *et al.*, *Environ Microbiol*. **11**,176-187 (2009).

10. F. Thevenieau, M. L. Fardeau, B. Ollivier, C. Joulian, S. Baena, *Extremophiles* **11**, 295-303 (2007).

11. S. Hattori, Y. Kamagata, S. Hanada, H. Shoun, *Int J Syst Evol Microbiol*. **50**,1601-1609 (2000).

12. A. I. Slobodkin *et al., Curr. Microbiol.* **39**, 99-102 (1999).

13. M. B. Salinas *et al., Int J Syst Evol Microbiol*. **54**, 645-649 (2004).

14. K. O. Stetter *et al., Nature* **365**, 743-745 (1993).

15. C. Jeanthon, O. Nercessian, E. Corre, A. Grabowski-Lux, in *Petroleum Microbiology*, B. Ollivier, M. Magot, Eds. (ASM Press, Washington, D.C., 2005) pp. 55-70.

16. L. Cheng *et al*., *Int J Syst Evol Microbiol.* **57**, 2964-2969 (2007).

17. M. W. Maune, R. S. Tanner, *Abst. Ann. Meet. Am. Soc. Microbiol.*, I-029, p. 321 (2008).

18. R. J. Menes L. Muxí, *Int J Syst Evol Microbiol*. **52**, 157-164 (2002).

19. H. L. Drake, in *Acetogenesis*, H. L. Drake, Ed. (Chapman & Hall, New York, 1994), pp. 3-60.

20. R. F. Mueller, P. H. Nielsen, *Appl. Environ. Microbiol*. 62, 3083-3087 (1996).

21. I. A. Davidova, J. M. Suflita, in *Methods in Enzymology*, J. Leadbetter, Ed. (Elsevier, San Diego, CA, 2005), vol. 397, pp. 17-34.

22. L. M. Gieg, J. M. Suflita, in *Petroleum Microbiology*, B. Ollivier, M. Magot, Eds. (ASM Press, Washington, D.C., 2005), pp. 337-356.

23. O. Kniemeyer *et al*., *Nature* **449,** 898-901 (2007).

24. H. Wilkes *et al., Arch Microbiol* **177**, 235-243 (2002).

25. S. L. Caldwell *et al*., *Environ. Sci Technol*. **42**, 6791-6799 (2008).

26. K. G. Lloyd, L. Lapham, A. Teske, *Appl. Environ. Microbiol.* **72**, 7218–7230 (2006).

27. B. Hedges, L. McVeigh, "The role of acetate in CO_2 corrosion: the double whammy" (Paper #21, NACE Corrosion 99,1999).

28. Support from the National Science Foundation (Award # 0647712) and ConocoPhillips is gratefully acknowledged. We particularly appreciate the assistance of Dr. Gary Jenneman and the Alaska Business Unit personnel of ConocoPhillips for organizing, collecting and shipping samples as well as for the use of ANS laboratory

facilities. The conclusions expressed in this paper are those of the authors and not necessarily shared by ConocoPhillips.

Supporting Online Material

Materials and Methods

Tables S1 to S3

References

Table 1. Metabolites associated with the anaerobic biodegradation of C_1-C_4 in Alaskan North Slope (ANS) oil field samples. Alkylsuccinates were detected in processing facility and production well samples (total=6), but not in seawater or in a pipeline transporting seawater, suggesting the anaerobic oxidation of the parent compounds methane, ethane, propane or butane. Concentrations of metabolites were in the μM range. Downstream metabolites resulting from the predicted carbon skeleton rearrangement and subsequent decarboxylation of the alkylsuccinate were also found in ANS samples. For *n*-alkanes C_3 or greater, a terminal and subterminal addition of fumarate (denoted with *) are possible, resulting in two possible downstream metabolites (branched or straight chain).

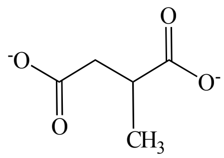
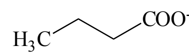
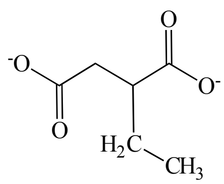
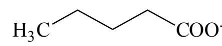
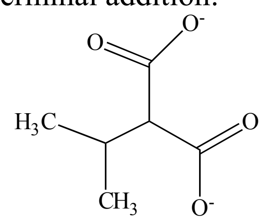
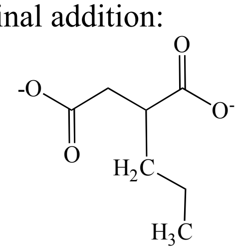
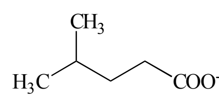
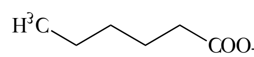
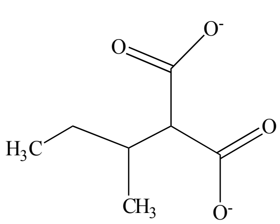
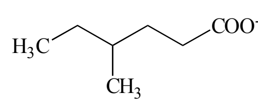
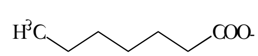
Parent compound	Fumarate addition metabolite	Fumarate addition metabolite concentration detected (µM)	Downstream metabolite (rearrangement)
Methane (CH_4)	 Methylsuccinate	2.08 ± 1.10	 Butanoic acid
Ethane (C_2H_6)	 Ethylsuccinate	1.77 ± 1.54	 Pentanoic acid
Propane (C_3H_8)	subterminal addition:  terminal addition:  Propylsuccinate*	2.18 ± 0.20	 4-Methylpentanoic acid  Hexanoic acid
Butane (C_4H_{10})	 Butylsuccinate*	0.76 ± 0.11	 4-Methylhexanoic acid  Heptanoic acid

Figure legends

Figure 1. Relative abundances of sequences from five bacterial 16S rRNA gene libraries. The RDP Project Classifier tool (http://rdp.cme.msu.edu/classifier/classifier.jsp) was used to assign representative sequences (97% similarity) to the higher-level taxonomic groups shown, except for sequences affiliated with *Thermovirga lienii* (referred to as "Synergistes" in this figure), which currently are classified under "Clostridia", Incertae sedis XV. N represents the total number of sequences in a library, after exclusion of chimeric sequences. Origin of samples: production wells 2P and 2L (sampled at wellhead), outflow from the 1^{st} stage separator (PS) and coalescer (CO) units in a central facility, and the "pig envelope" (e.g. the scraped inner surface of the pipeline) of a pipeline transporting treated seawater (SW) from the Arctic Ocean to the central facility. Table S1A contains the accession number of the closest match, affiliation, and the relative abundance (as a percentage of the total library) of each representative sequence. Temperatures at the 2P wellhead for 9 months prior to sampling ranged from 36-52°C, avg. 48°C ±1.7 (1 SD); from the 2L wellhead 33-52°C, avg. 49°C ±2.1 (1 SD). Temperature ranges inside the central facility were maintained at 50-55°C (1^{st} stage separator), 67-82°C (2^{nd} stage separator) and 59-78°C (coalescer).

Figure 2. Distribution of bacterial sequences from 2L, PS, and CO illustrating the number of sequences from taxa found in all three libraries ("2L, PS, CO: Core taxa"), 2 libraries ("PS and 2L", "CO and 2L", "PS and CO") or unique to one sample.

Figure 3. Relative abundances of sequences from three archaeal 16S rRNA gene libraries. The RDP Project Classifier tool (http://rdp.cme.msu.edu/classifier/classifier.jsp) was used to assign representative sequences (97% similarity) to the higher-level taxonomic groups

shown. N represents the total number of sequences in a library, after exclusion of chimeric sequences. Origin of samples: production well 2P (sampled at wellhead), outflow from the 1st stage separator (PS) and coalescer (CO) units in a central facility. Table S1B contains the accession number of the closest match, affiliation, and the relative abundance (as a percentage of the total library) of each representative sequence.

Figure 1.

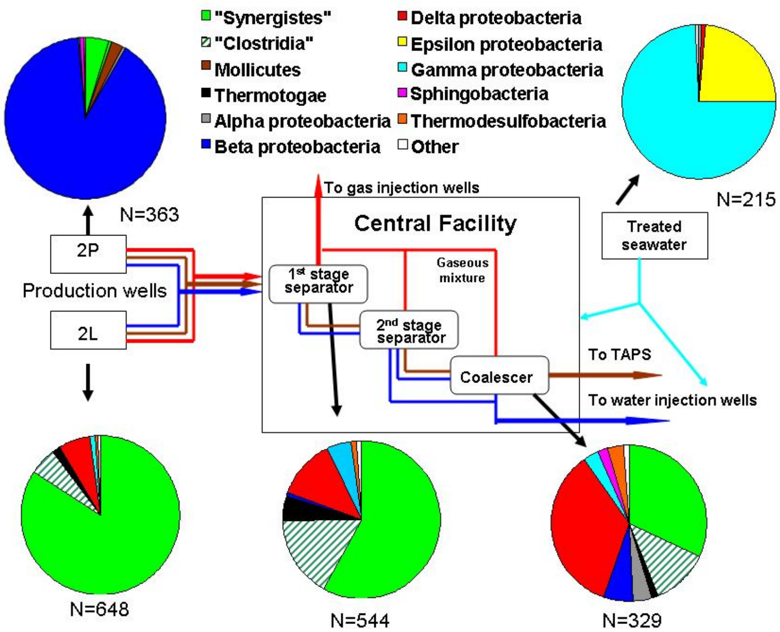
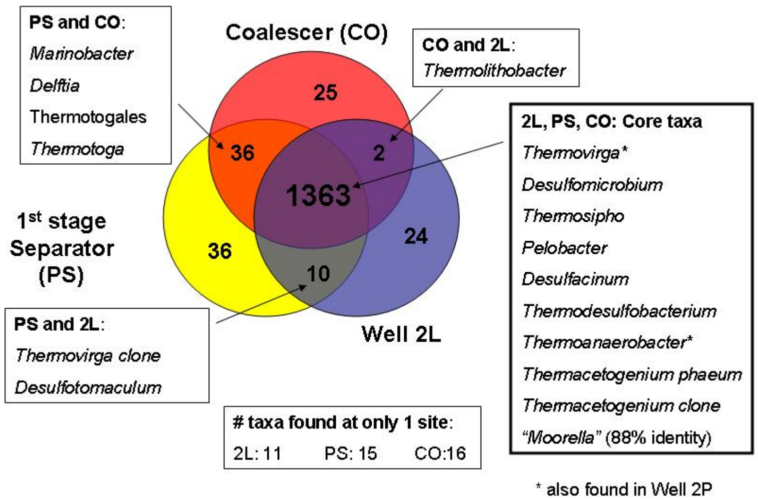
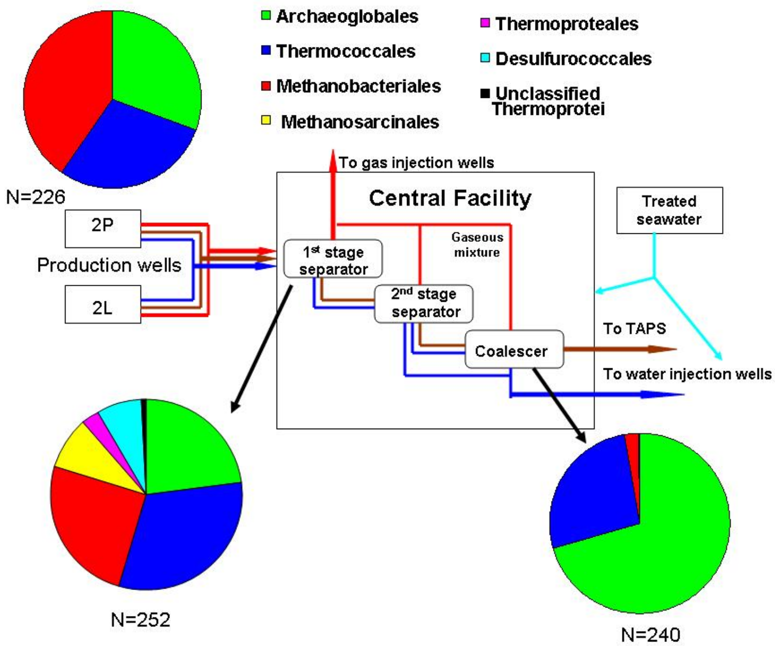

"Synergistes"
"Clostridia"
Mollicutes
Thermotogae
Alpha proteobacteria
Beta proteobacteria
Delta proteobacteria
Epsilon proteobacteria
Gamma proteobacteria
Sphingobacteria
Thermodesulfobacteria
Other
N=363
N=215
To gas injection wells
Central Facility
2P
Production wells
2L
1st stage separator
2nd stage separator
Gaseous mixture
Coalescer
Treated seawater
To TAPS
To water injection wells
N=648
N=544
N=329

Fig. 2.



PS and CO:
Marinobacter
Delftia
Thermotogales
Thermotoga
Coalescer (CO)
CO and 2L:
Thermolithobacter
25
36
2
1363
1st stage Separator (PS)
36
24
10
2L, PS, CO: Core taxa
Thermovirga*
Desulfomicrobium
Thermosipho
Pelobacter
Desulfacinum
Thermodesulfobacterium
Thermoanaerobacter*
Thermacetogenium phaeum
Thermacetogenium clone
"Moorella" (88% identity)
PS and 2L:
Thermovirga clone
Desulfotomaculum
Well 2L
taxa found at only 1 site:
2L: 11 PS: 15 CO:16
* also found in Well 2P

Fig. 3.


Archaeoglobales
Thermococcales
Methanobacteriales
Methanosarcinales
Thermoproteales
Desulfurococcales
Unclassified Thermoprotei
N=226
2P
Production wells
2L
To gas injection wells
Central Facility
1st stage separator
Gaseous mixture
2nd stage separator
Coalescer
Treated seawater
To TAPS
To water injection wells
N=252
N=240

Supporting online material

Materials and Methods

Molecular analysis: Samples from two production wells (2P, 2L), two locations in a central facility (CF, 1st stage separator [PS] and coalescer [CO]), and from a seawater line prior to exposure to oil (SW) were collected in 2006 from an oil field complex on the North Slope of Alaska. The seawater line sample consisted of fluids and solids from a pigging operation, whereas the other samples were fluids. Two samples (PS and SW), 150 mL each) were filtered (0.45 µm) and preserved in the field by the addition of DNAzol® Direct (Molecular Research Center, Inc., Cincinnati, OH) to the filter then extracted at OU using a bead-beating protocol (UltraClean™ Mega Soil DNA Isolation, MO BIO Laboratories, Inc., Carlsbad, CA). The remaining samples (20 mL) were first concentrated by ethanol precipitation and the pellet resuspended in PCR-grade water before extraction using a bead-beating protocol (PowerSoil™ DNA Isolation Kit, MO BIO). 16S rRNA primers for eubacteria (16S/18S rRNA PCR Library Creation, http://my.jgi.doe.gov/general/index.html, and 27F and 1492R for the seawater line *(1)*, ARC333F and 958R for archaea *(1)*, *dsrAB* (dsr1F, dsr4R, *2*), and *mcrA* primers (ME1 and ME2, *3*) were used to obtain PCR products to create clone libraries (5 eubacterial 16S, 3 archaeal 16S, 4 *dsrAB*, 1 *mcrA*) using the TOPO® TA Cloning Kit (Invitrogen Corp., Carlsbad, CA). A sixth, duplicate bacterial 16S library using primers 27F and 1492R *(1)* was created from one production well sample (2P) to compare the effect of possible primer bias. Sequencing of the libraries was performed by the DOE Joint Genome Institute (Lawrence Livermore Laboratory, Walnut Creek, CA). Results for the *dsrAB* and *mcrA* libraries are briefly referred to in this work and will be reported in detail later (manuscript in preparation). The duplicate bacterial 16S library created with primers 27F and 1492R from sample 2P gave the same dominant *Petrobacter* sequence (87.7% of total sequences) and low sequence diversity as did the library created with 27F and 1391 (90.1% of total sequences were *Petrobacter*, see tables S1A and S2) and will not be discussed further.

Primer binding sites were identified using the "Motifs" function in Sequencher (version 4.7, Gene Codes, Ann Arbor, MI) as a guide to trim the sequences to homologous regions, approximately 1250 bp for bacterial 16S rRNA gene sequences and 600 bp for archaea. The sequences in the clone libraries were aligned using the greengenes NAST-aligner *(4)* and examined for chimeric sequences using the Bellerophon program *(5)* available through the greengenes website (version 3, http://greengenes.lbl.gov/cgi-bin/nph-bel3_interface.cgi). Potential chimeric sequences identified by Bellerophon were further examined by Pintail *(6)* and comparing separate regions of the sequences by BLASTN *(7)*. Distance matrices *(8*, greengenes, "Create distance matrix", http://greengenes.lbl.gov/cgi-bin/nph-distance_matrix.cgi) were created from each library after it had been purged of chimeras. The Lane mask filter *(9)* was applied to limit distance matrix calculations to conserved portions of the aligned sequences. DOTUR *(10)* was used to create the distance matrix to produce OTUs at the 97% level of similarity and calculate the Chao and ACE estimates of species richness and Shannon-Weaver and Simpson measures of diversity reported in table S2. % library coverage at the 97% level of similarity was estimated by the method of Good *(11)*. One representative sequence was chosen from each OTU, its taxonomic affiliation determined by the RDP Classifier *(12)* and closest match to sequences in the GenBank database by

BLASTN (*7*; tables S1A and S1B). Distance matrices were constructed from the pooled representative sequences originating from all libraries of the same type (e.g. 5 bacterial libraries pooled into one, 3 archaeal libraries pooled into one) and DOTUR applied to produce pooled-sample OTUs at the 97% level of similarity. Correct assignment of pooled-sample OTU membership for each individual representative sequence within a pooled-sample OTU was confirmed by inspection of the taxonomic affiliation and BLASTN matches previously determined for the representative sequence.

Representative bacterial sequences were deposited in GenBank under accession numbers FJ269280-FJ269403; representative archaeal sequences were assigned accession numbers FJ446497-FJ446523.

Enrichments and isolation

General heterotrophs were enumerated in anaerobic half-strength tryptic soy broth (Becton, Dickinson and Co.) plus 1% NaCl under a N2:CO2 atmosphere in a MPN assay *(13)*. Hydrogen oxidizers were enumerated in an anaerobic, reduced basal medium under a H2:CO2 atmosphere *(13)*. SRB were enumerated in a medium designed for rapid quantitation *(13)*.

Metabolic profiling: Fluids (1 L) from the North Slope oil field were collected from central facility (PS), production wells (2L, 2P, 2T), a pipeline carrying fluids and gas to the central facility (2U), a water reinjection well (2K) and seawater lines (biocide-treated seawater "TS", and fluids and solids scraped from the inner surface of the pipeline carrying treated seawater "SW") and immediately preserved in the field with 50% HCl (pH < 2) for metabolite analysis. Samples TS and SW were used to provide background values as they do not contain hydrocarbons. Acidified samples were kept at room temperature until they were extracted with ethyl acetate, dried over anhydrous Na_2SO_4, and concentrated by rotary evaporation and under a stream of N_2. Concentrated extracts were derivatized with *N,O*-bis(trimethylsilyl)trifluoroacetamide (BSTFA) (Pierce Chemical Co., Rockford, IL) to add trimethylsilyl groups for analysis by gas chromatography-mass spectrometry (GC-MS). Derivatized components were separated on a HP-5ms capillary column (30 m x 0.25 mm i.d., J&W Scientific, Folsom, CA) with a starting oven temperature of 45°C (held 5 min) increasing at 4°C/min to 270°C (held 10 min) before mass spectral analysis. Metabolite identifications were made by comparison with the GC-MS features of authentic standards or with previously reported MS profiles *(14-16)*.

Table S1A. Sequence similarity and taxonomic relationships of bacterial representative small subunit partial rRNA gene sequences (OTUs at 97% similarity)

OTU	% of total	Accession number	Most similar sequences* (accession no.)	%	Class**	Source
Well 2P: total # sequences =363, primers 27F, 1391R. 97% OTUs from DOTUR						
2P327SHNG718	90.1	FJ469286	*Petrobacter* sp. NFC7-F8 (EU250943)	99	β proteo	50°C compost
			Petrobacter sp. DM-3 (DQ539621)	99		Dagang oil field
2P17SHNG539	4.7	FJ469280	*Thermovirga lienii* Cas60314 (DQ071273)	99	Synergistes	North Sea oil well
2P9SHNG554	2.5	FJ469288	Uncultured clone CK06-06_Mud_MAS1B-28 (AB369171)	99	Mollicutes	Offshore drilling mud fluid
2P3SHNG611	0.8	FJ469287	Uncultured clone B5_B4 (EF025213)	99	Sphingobacteria	Turkey intestine
2P2SHNG411	0.6	FJ469285	*Thermoanaerobacter pseudethanolicus* ATCC 33223 (CP000924)	99	"Clostridia"	Octopus Springs
2P2SHNG385	0.6	FJ469284	*Bradyrhizobium* sp. JR016 (EF221629)	99	α proteo	Root nodule
2P1SHNG397	0.3	FJ469281	Uncultured Thermotogales clone bh459.f1.4.b07 (AM184116)	99	Thermotogae	Low-temp enrichment degrading polychlorinated biphenyls
2P1SHNG452	0.3	FJ469282	*Stenotrophomonas* sp. ROi7 (EF219038)	99	γ proteo	Reverse osmosis membrane
2P1SHNG731	0.3	FJ469283	*Ralstonia pickettii* 12J (CP001069)	99	β proteo	
Well 2L: total # sequences = 648, primers 27F, 1391R. 97% OTU from DOTUR.						
2L474SGXO1136	73.1	FJ469308	*Thermovirga lienii* Cas60314 (DQ071273)	99	Synergistetes	North Sea oil well
2L65SGXO482	10.0	FJ469310	*Thermovirga lienii* Cas60314 (DQ071273)	99	Synergistes	North Sea oil well
2L27SGXO638	4.2	FJ469303	*Desulfomicrobium thermophilum* P6.2 (AY464939)	98	δ proteo	Hot spring in Colombia
2L14SGXO613	2.2	FJ469289	*Thermoanaerobacter pseudethanolicus* ATCC 33223 (CP000924)	99	"Clostridia"	Octopus Springs
2L9SGXO552	1.4	FJ469316	*Thermosipho africanus* (DQ647057)	99	Thermotogae	Shallow hydrothermal system
			Thermosipho sp. TBA5 AF231727	99		North Sea oil field
2L7SGXO640	1.1	FJ469315	*Desulfacinum subterraneum* (AF385080)	97	δ proteo	High temp Vietnam oil field
2L6SGXO418	0.9	FJ469313	Uncultured *Thermovirga* sp. clone TCB169x (DQ647105)	95	Synergistetes	North Sea oil well
2L6SGXO1151	0.9	FJ469311	Uncultured organism clone ctg_NISA224 (DQ396164)	95	γ proteo	Deep-sea octacoral
			Shewanella sp. IS5	95		Diseased larval rock

			(AY967729)			lobster cultures
2L6SGXO579	0.9	FJ469314	Uncultured clone Niigata-10 (AB243821)	99	δ proteo	Niigata (Japan) oil well
			Pelobacter carbinolicus (CP000142)	97		
2L6SGXO407	0.9	FJ469312	*Thermacetogenium phaeum* strain PBT (AB020336)	98	"Clostridia"	Thermophilic anaerobic methanogenic reactor
2L5SGXO560	0.8	FJ469309	*Clostridium* sp. C9 (EU862317)	99	"Clostridia"	Off-shore oil well, India
2L3SGXO643	0.5	FJ469306	*Thermodesulfobacterium commune* DSM 2178 (AF418169)	99	Thermodesulfobacteria	Yellowstone thermal spring
2L3SGXO888	0.5	FJ469307	Uncultured clone Niigata-15 (AB243826)	99	"Clostridia"	Niigata (Japan) oil well
2L2SGXO601	0.3	FJ469304	*Dehalococcoides* sp. CBDB1 (AF230641)	99	Chloroflexi Dehalococcoides	Methanogenic enrichment from Saale river sediment
2L2SGXO622	0.3	FJ469305	Gram-positive thermophile strain ODP159-02 (AY704384)	94	"Clostridia"	Ocean ridge flank crustal fluid
2L1SGXO770	0.2	FJ469299	*Sulfurospirillum* sp. NO3A (AY135396)	99	ε proteo	Coleville (Canada) oil field
2L1SGXO762	0.2	FJ469298	*Thermolithobacter thermoautotrophicus* KA2b (AF282254)	98	Thermolithobacteria	Yellowstone Calcite Springs
2L1SGXO814	0.2	FJ469301	Uncultured *Natronoanaerobium* sp. clone SHBZ503 (EU639010)	88	"Clostridia"	Thermophilic microbial fuel cell
			Moorella thermoacetica ATCC 39073 (CP000232)	88		Horse manure
2L1SGXO566	0.2	FJ469295	*Flexistipes* sp. vp180 (AF220344)	98	Deferribacteres	High temperature oil reservoir
2L1SGXO817	0.2	FJ469302	*Desulfotomaculum thermocisternum* strain ST90 (U33455)	99	"Clostridia"	Hot North Sea oil reservoir
2L1SGXO1041	0.2	FJ469292	*Pseudomonas stutzeri* strain 24a97 (AJ312172)	99	γ proteo	Soil beneath filling station
2L1 SGXO442	0.2	FJ469293	Uncultured bacterium clone PL-25B8 (AY570610)	99	"Clostridia"	Low-temperature biodegraded Canadian oil reservoir
			Acetobacterium carbinolicum (AY744449)	99		
2L1SGXO751	0.2	FJ469297	Uncultured bacterium clone PL-38B5 (AY570590)	99	"Clostridia" (Anaerovorax , 100%)	Low-temperature biodegraded Canadian oil reservoir
2L1SGXO459	0.2	FJ469294	*Thermosipho africanus* strain Ob7 (DQ647057)	99	Thermotogae	
			Thermosipho sp. TBA5 (AF231727)	99		North Sea oil field
2L1SGXO697	0.2	FJ469296	*Thermovirga lienii* Cas60314 (DQ071273)	97	Synergistetes	North Sea oil well
2L1SGXO811	0.2	FJ469300	Uncultured *Thermacetogenium* sp. clone B11_otu13 (DQ097678)	97	“Clostridia”	High temperature Dagang oil field (China)
2L1SGXO1021	0.2	FJ469290	*Thermosipho geolei* (AJ272022)	99	Thermotogae	Siberian oil reservoir
2L1SGXO1038	0.2	FJ469291	Uncultured Spirochaetaceae clone	99		North Sea oil field

			TCB129x (DQ647164)	99		Congo oil field
			Spirochaeta sp. MET-E (AY800103)			
1st stage separator: total= 544 sequences, primers 27F, 1391 R. 97% OTUs from DOTUR.						
PS313SGXI1055	57.7	FJ469337	*Thermovirga lienii* Cas60314 (DQ071273)	99	Synergistetes	North Sea oil well
PS74SGXI1247	13.6	FJ469348	*Thermoanaerobacter pseudethanolicus* ATCC 33223 (CP000924)	99	"Clostridia"	Octopus Springs
			Thermoanaerobacter strain X514(CP000923)	99		Colorado deep subsurface, iron-reducing
PS39SGXI1143	7.2	FJ469338	*Desulfomicrobium thermophilum* P6.2 (AY464939)	99	δ proteo	Hot spring in Colombia
PS20SGXI1921	3.7	FJ469333	Uncultured bacterium clone Niigata-10 (AB243821)	99	δ proteo	Niigata (Japan) oil well
			Pelobacter carbinolicus DSM 2380 (CP000142)	97		
PS10SGXI1101	1.8	FJ469317	*Thermosipho africanus* (DQ647057)	99	Thermotogae	Hot North sea oil field
PS9SGXI1270	1.7	FJ469351	*Thermacetogenium phaeum* (AB020336)	99	"Clostridia"	Thermophilic anaerobic methanogenic reactor
PS8SGXI1894	1.5	FJ469350	*Halomonas meridiana* strain aa-9 (EU652041)	99	γ proteo	Ocean sediment
PS8SGXI1020	1.5	FJ469349	*Halomonas* sp. A-07 (AY347310)	99	γ proteo	Tanzania soda lakes
PS6SGXI1029	1.1	FJ469345	*Thermodesulfobacterium commune* DSM 2178 (AF418169)	99	Thermodesulfo-bacteria	Yellowstone thermal spring
PS6SGXI904	1.1	FJ469347	*Thermotoga petrophila* RKU-1 (CP000702)	99	Thermotogae	Kubiki oil reservoir, Niigata, Japan
PS6SGXI1183	1.1	FJ469346	*Desulfacinum subterraneum* (AF385080)	97	δ proteo	High temp Vietnam oil field
PS5SGXI1002	0.9	FJ469344	Uncultured bacterium clone S25_271 (EF573927)	99	γ proteo	Costa Rica island
			Marinobacter bacchus strain FB3 (DQ282120)	99		Evaporation pond of wine wastewater
PS4SGXI1172	0.7	FJ469341	Uncultured ***Natronoanaerobium sp.*** clone SHBZ503 (EU639010)	87	"Clostridia"	Thermophilic microbial fuel cell
			Moorella thermoacetica AMP *(*AY884087)	87		Methanogenic sludge
PS4SGXI910	0.7	FJ469343	*Delftia acidovorans* SPH-1 (CP000884)	99	β proteo	Sewage treatment plant
PS4SGXI1530	0.7	FJ469342	*Thermotoga naphthophila* RKU-10 (AB027017)	87	Thermotogae	Kubiki oil reservoir, Niigata, Japan
PS4SGXI1065	0.6	FJ469339	*Marinobacter hydrocarbonoclasticus* MARC4F (DQ768638)	99	γ proteo	Middle Atlantic Ridge Sediment
PS3SGXI1245	0.6	FJ469340	Uncultured clone MAT-CR-H3-B03 (EU245152)	85	unclassified	Hypersaline microbial mat, P.R.
PS2SGXI1098	0.4	FJ469335	*Petrotoga siberica* strain SL25T (AJ311702)	99	Thermotogae	Siberian oil reservoir
PS2SGXI1003	0.4	FJ469334	*Desulfotomaculum thermocisternum* (U33455)	99		Hot North Sea oil reservoir

PS2SGXI1381	0.4	FJ469336	Thermotogales TBF19.5.1 (EU980631)	99	Thermotogae	North Sea oil production fluid
PS1SGXI1064	0.2	FJ469318	*Thermosipho geolei.* DSM 13256 (AJ272022)	98	Thermotogae	Siberian oil reservoir
PS1SGXI1111	0.2	FJ469319	Uncultured bacterium clone Zplanct13 (EF602474)	93	unclassified	Zodletone Spring source sediments
PS1SGXI1133	0.2	FJ469320	Uncultured *Sulfurospirillum* sp. clone LA4-B52N (AF513952)	93	ε proteo	Hawaiian lake water
			Sulfurospirillum carboxydovorans (AY740528)	93		North Sea sediment
PS1SGXI1244	0.2	FJ469321	*Geotoga aestuarianus* strain T3B (AF509468)	99	Thermotogae	Karst sink hole thiosulfate-reducer
PS1SGXI1249	0.2	FJ469322	*Burkholderia multivorans* strain LMG 13010^T (Y18703)	99	β proteo	Cystic fibrosis patient
PS1SGXI1260	0.2	FJ469323	Uncultured bacterium clone: HDBW-WB60 (AB237723)	99	"Clostridia"	Deep subsurface groundwater
PS1SGXI1272	0.2	FJ469324	*Thermovirga lienii* Cas60314 (DQ071273)	97	Synergistetes	North Sea oil well
PS1SGXI1281	0.2	FJ469325	Uncultured *Thermovirga* sp. clone TCB8y	97	Synergistetes	North Sea produced water
PS1SGXI1300	0.2	FJ469326	*Marinobacterium* sp. IC961 strain IC961	99	γ proteo	Carbazole-utilizing bacterium
PS1SGXI1313	0.2	FJ469327	Uncultured bacterium a2b00 (AF419657)	92	unclassifed	Hydrothermal sediments in the Guaymas Basin
PS1SGXI1413	0.2	FJ469328	Uncultured *Thermacetogenium* sp. clone B11_otu13 (DQ097678)	98	"Clostridia"	High temperature Dagang oil field (China)
PS1SGXI1848	0.2	FJ469329	*Pseudomonas putida* W619 (CP000949)	100	γ proteo	Japan:Shizuoka, Sagara oil field
			Pseudomonas sp. OCR2 (AB240201)	99		
PS1SGXI1964	0.2	FJ469330	Uncultured *Thermovirga* sp. clone TCB169x (DQ647105)	96	Synergistetes	High temp North Sea oil field
PS1SGXI1984	0.2	FJ469331	Uncultured bacterium gene (AB195893)	96	Bacteroidetes	Anaerobic sludge
PS1SGXI995	0.2	FJ469332	*Desulfotignum balticum* DSM 7044 (AF418176)	99	δ proteo	Marine coastal sediment, Baltic Sea
Coalescer: total # sequences = 329. primers 27F, 1391 R.97% OTUs from DOTUR.						
CO105SHNF404	31.9	FJ469352	*Thermovirga lienii* Cas60314 (DQ071273)	99	Synergistetes	High temp North Sea oil field
			Uncultured *Thermovirga* sp. clone TCB8y (DQ647105)	99		High temp North Sea oil field
CO60SHNF483	18.2	FJ469383	*Desulfocaldus* sp. Hobo (EF442977)	99	δ proteo	Not specified
			Desulfomicrobium thermophilum P6.2 (AY464939)	99	δ proteo	Hot spring in Colombia
CO46SHNF563	14.0	FJ469378	Uncultured bacterium clone: Niigata-10 (AB243821)	99	δ proteo	Niigata (Japan) oil well
			Pelobacter carbinolicus DSM 2380 (CP000142)	97		

CO26SHNF710	7.9	FJ469371	*Thermoanaerobacter pseudethanolicus* ATCC 33223(CP000924)	100	"Clostridia"	Octopus Springs
CO12SHNF562	3.6	FJ469354	Uncultured bacterium clone cc187 (DQ057384)	100	β proteo	Chicken intestine
			Beta proteobacterium B7 AF035053	98		Drinking water system
CO11SHNF516	3.3	FJ469353	*Thermodesulfobacterium commune* DSM 2178 (AF418169)	99	Thermodesulfobacteria	Yellowstone thermal spring
CO7SHNF526	2.1	FJ469386	*Stenotrophomonas maltophilia* strain DN1.1 (EU034540)	99	γ proteo	Not specified
			Uncultured bacterium clone Ana10UA-2 (EU499720)	99	γ proteo	Freshwater sediment
CO6SHNF422	1.8	FJ469384	*Bradyrhizobium japonicum* strain SEMIA 6164 (AY904765)	99	α proteo	Acacia root nodule
CO6SHNF588	1.8	FJ469385	*Desulfacinum subterraneum* (AF385080)	98	δ proteo	High temp Vietnam oil field
CO5SHNF732	1.5	FJ469382	*Thermacetogenium phaeum* strain PBT (AB020336)	99	"Clostridia"	Thermophilic anaerobic methanogenic reactor
CO5SHNF565	1.5	FJ469380	*Ralstonia pickettii* 12J (CP001069)	100	β proteo	
			Uncultured bacterium clone P7X3b4E02 (EU491068)	100	β proteo	Seafloor lavas from the Loi'hi Seamount South Rift X3
CO5SHNF607	1.5	FJ469381	Uncultured clone B5_B4 (EF025213)	100	Sphingobacteria	Turkey intestine
			Sediminibacterium salmoneum (EF407879)	96		Eutrophic reservoir
CO4SHNF461	1.2	FJ469379	*Mesorhizobium plurifarium*,strain LMG 10056 (Y14161)	99	α proteo	Tropical tree
CO3SHNF446	0.9	FJ469375	Uncultured *Natronoanaerobium* sp. clone SHBZ503 (EU639010)	88	"Clostridia"	Microbial fuel cell
			Moorella thermoacetica ATCC 39073 (CP000232)	88		Horse manure
CO3SHNF573	0.9	FJ469376	Beta proteobacterium A1040 (AF236008)	99	β proteo	Not specified
			Beta proteobacterium MB7 (AB013409)	99		Soil isolate degrading aliphatic polyesters
CO3SHNF586	0.9	FJ469377	*Hyphomicrobium* sp. P2 (AF148858)	99	α proteo	Portuguese soil
CO2SHNF510	0.6	FJ469373	*Thermosipho africanus* (DQ647057)	99	Thermotogae	Shallow hydrothermal system
			Thermosipho sp. TBA5 AF231727	99		North sea oil field
CO2SHNF508	0.6	FJ469372	*Solemya velum* symbiont (M90415)	99	γ proteo	Sulfur-oxidizing mollusk symbiont
CO2SHNF712	0.6	FJ469374	*Spirochaeta thermophila* (X62809)	98	Spirochaetes	Kuril Island hot springs
CO1SHNF389	0.3	FJ469355	Uncultured *Hydrogenothermus* sp. clone OPPB154 (AY861874)	99	Aquificales	Yellowstone Obsidian Pool
			Aquificales bacterium	99		Yellowstone Calcite

			YNP-SS1 (AF507961)			Springs
CO1SHNF407	0.3	FJ469356	*Desulfomicrobium norvegicum* strain DSM 1741T (AJ277897)	99	δ proteo	Oslo Harbour water
CO1SHNF410	0.3	FJ469357	*Marinobacter bacchus* strain FB3 (DQ282120)	99	γ proteo	Wine wastewater
CO1SHNF443	0.3	FJ469358	*Thermotoga elfii* strain SM-2 (EU276416)	99	Thermotogae	Not specified Oil-production water
CO1SHFN484	0.3	FJ469359	Uncultured bacterium clone rRNA082 (AY958855)	98	Sphingobacteria	Human vaginal epithelium
			Solibium sp. I-32 (AM990455)	97		Ultra pure water
CO1SHNF528	0.3	FJ469360	Uncultured clone B5_F26 (EF025264)	99	Sphingobacteria	Turkey intestine
			Flavobacteria bacterium KF030 (AB269814)	94		Freshwater lake
CO1SHNF536	0.3	FJ469361	Uncultured bacterium clone PS18 (DQ984666)	93	"Clostridia"	Sulfate-reducing LCFA enrichment
			Syntrophomonas palmitatica (AB274040)	93		Methanogenic sludge
CO1SHFN544	0.3	FJ469362	Uncultured *Thermacetogenium* sp. clone B11_otu13 (DQ097678)	98	"Clostridia"	High temperature Dagang oil field (China)
CO1SHNF610	0.3	FJ469363	*Desulfovibrio aespoeensis* clone Aspo3 (EU680957)	98	δ proteo	Aespoe hard rock
			Desulfovibrio aespoeensis isolate Aspo2 (X95230)			
CO1SHNF612	0.3	FJ469364	*Thermotoga petrophila* RKU-1 (AJ872269)	100	Thermotogae	Kubiki oil reservoir
CO1SHNF615	0.3	FJ469365	Uncultured Termite group 1 bacterium clone HAVOmat14 (EF032762)	99	candidate division TG1	Cyanobacterial mat in Hawaiian lava cave
CO1SHNF622	0.3	FJ469366	Thermotogales bacterium 2SM-2 (EU276414)	100	Thermotogae	Oil-production water
CO1SHNF639	0.3	FJ469367	*Thermolithobacter thermoautotrophicus* clone KA2b (AF282254)	99	Thermolithobacteria	Yellowstone Calcite Springs
CO1SHNF644	0.3	FJ469368	Thermoanaerobacteriaceae clone EV818FW062101BH4MD48 (DQ079638)	95	"Clostridia"	Terrestrial subsurface fluid-filled fracture
			Moorella thermoacetica strain AMP (AY884087)	95		Methanogenic sludge
CO1SHNF669	0.3	FJ469369	Uncultured *Anaerovorax* sp. clone C14B-1H (EU073780)	97	"Clostridia"	Coal enrichment culture
			Clostridiaceae bacterium FH042 (AB298771)	96		Anaerobic sludge of a methanogenic reactor
CO1SHNF695	0.3	FJ469370	*Desulfovibrio* sp. X (EF442979)	99	δ proteo	Not specified
			Desulfovibrio zosterae (Y18049)	95		Roots of seagrass (*Zostera marina*)
Seawater pig envelope. Total # sequences = 215. primers 27F, 1492R. 97% OTUs from DOTUR						
SW76FGIT720	35.3	FJ469401	*Pseudomonas* sp. HZ06 (AY690706)	99	γ proteo	Rhizosphere soil of salt marshes
SW55FGIT591	25.6	FJ469399	*Pseudomonas stutzeri* strain aa-28 (EU652047)	99	γ proteo	Ocean sediment
SW35FGIT483	16.3	FJ469395	Uncultured proteobacterium clone	97	ε proteo *Sulfurimonas*	Guaymas Basin hydrothermal vent

			B01R008 (AY197379)			sediments
SW12FGIT423	5.6	FJ469387	*Pseudomonas* sp. EP27 (AM403529)	98	γ proteo	Deep-sea sediments
SW9FGIT664	4.2	FJ469403	Uncultured *Arcobacter* sp. clone DS172 (DQ234254)	98	ε proteo *Arcobacter*	Mangrove
SW7FGIT497	3.3	FJ469402	Uncultured bacterium clone W26 (AY770966)	98	γ proteo *Pseudomonas*	Water injection well of Dagang oilfield
SW5FGIT667	2.3	FJ469400	Uncultured *Pseudomonas* sp. clone Lupin-1130m-2-MDA-pse3 (EF205269)	98	γ proteo	Lupin gold mine fracture water
SW3FGIT405	1.4	FJ469396	*Pseudomonas marincola* (AB301071)	96	γ proteo	Deep-sea brittle star
SW3FGIT554	1.4	FJ469397	Uncultured alpha proteobacterium clone 131582 (AY922182)	97	ε proteo *Arcobacter*	Grey whale bone, Pacific Ocean, depth 1674 meters
SW3FGIT592	1.4	FJ469398	Uncultured epsilon proteobacterium clone: NKB11 (AB013263)	96	ε proteo *Sulfurimonas*	Nankai Trough sediments
SW1FGIT389	0.5	FJ469388	"Gamma" proteobacterium IR (AF521582)	99	γ proteo /unclassified	Not specified "Diversity of marine humics-oxidizing bacteria"
SW1FGIT424	0.5	FJ469389	Uncultured bacterium clone B8S-8 (EU652615)	88	δ proteo	Yellow Sea sediment
SW1FGIT462	0.5	FJ469390	*Phaeobacter arcticus* strain 20188 (DQ514304)	99	α proteo	Arctic marine sediment
SW1FGIT467	0.5	FJ469391	Uncultured bacterium clone GZKB9 (AJ853504)	97	ε proteo *Arcobacter*	Landfill leachate
SW1FGIT501	0.5	FJ469392	Uncultured delta proteobacterium clone d13 (AY062878)	98	δ proteo *Desulfuromonas*	Electrode surface
SW1FGIT563	0.5	FJ469393	Uncultured bacterium clone P9X2b3A09 (EU491225)	86	δ proteo /unclassified	Seafloor lavas
SW1FGIT660	0.5	FJ469394	Uncultured bacterium ARCTIC23_B_12 (EU795085)	99	Flavobacteria *Polaribacter*	Arctic

Table 1B. Sequence similarity and taxonomic relationships of archaeal representative small subunit partial rRNA gene sequences (OTUs at 97% similarity)

OTU	% of total	Accession number	Most similar sequences (accession no.)	%	Orders	Source
Well 2P: total # sequences =226, primers ARC333 and 958R, 97% OTUs from DOTUR						
2P66FGIP571	29.2	FJ446503	Archaeon enrichment culture clone PW5.2A (EU573152)	100	Thermococcales	Ekofisk oil field
			Thermococcus alcaliphilus DSM 10322 (AB055121)	100		
2P66FGIP425	29.2	FJ446502	Uncultured archaeon SSE_L4_E01(EU635901)	99	Archaeoglobales	Hot spring sediment
			Archaeoglobus fulgidus strain L3 (DQ374392)	97		
2P64FGIP517	28.3	FJ446501	*Methanothermobacter thermautotrophicus* strain JZTM (EF100758)	99	Methanobacteriales	Jiaozhou Bay sediment
			M. wolfeii strain KZ24a (DQ657904)	99		Dagang oil field
2P26FGIP436	11.5	FJ446499	Uncultured Methanobacteriaceae clone A1m_OTU 3 (DQ097668)	99	Methanobacteriales	Dagang oil field
			Methanothermobacter thermautotrophicus strain JZTM (EF100758)	96		Jiaozhou Bay sediment
2P2FGIP540	0.9	FJ446500	Archaeon enrichment culture clone PW30.6A (EU573155)	99	Archaeoglobales	Ekofisk oil field
			Archaeoglobus sp. NI85-A (AB175518)	99		Deep-sea hydrothermal vent chimney
2P1FGIP407	0.4	FJ446497	Uncultured archaeon SSE_L4_E01(EU635901)	98	Archaeoglobales	Hot spring sediment
			Archaeoglobus fulgidus strain L3 (DQ374392)	97		
2P1FGIP710	0.4	FJ446498	Uncultured clone QHO-A15 (DQ785496)	96	Methanobacteriales	High temperature oil field in China
			Methanobacterium sp. F (AB302952)	96		Rice paddy soil
1st stage separator: total # sequences =252 , primers ARC333 and 958R, 97% OTUs from DOTUR						
PS70SGXN402	27.8	FJ446513	*Thermococcus mexicalis* strain GY 869 (AY099181)	99	Thermococcales	Hydrothermal deep-sea vents
			Thermococcus sibiricus (AJ238992)	99		Siberian high-temperature oil reservoir
PS56SGXN497	22.2	FJ446512	*Archaeoglobus fulgidus* DSM 4304 (AE000782)	99	Archaeoglobales	
			Archaeoglobus fulgidus strain L3 (DQ374392)	99		
PS35SGXN478	13.9	FJ446511	Uncultured archaeon clone NAK1-a1 (DQ867048)	99	Methanobacteriales	High-temperature natural gas field
			Uncultured bacterium clone QHO-A27 (DQ785508)	99		High-temperature petroleum reservoir

			Methanothermobacter thermautotrophicus strain JZTM (EF100758)	96		Jiaozhou Bay sediment
PS29SGXN477	11.5	FJ446509	Uncultured archaeon clone NAK1-a1 (DQ867048)	99	Methanobacteriales	High-temperature natural gas field
			Methanothermobacter thermautotrophicus strain JZTM (EF100758)	99		Jiaozhou Bay sediment
PS22SGXN482	8.7	FJ446508	Methanogenic archaeon ZC-1 (DQ787474) "*Methermicoccus shengliensis*"	99	Methanosarcinales	Oil production water
PS11SGXN711	4.4	FJ446504	Uncultured Desulfurococcales YNP_SSp_A61 (DQ243776)	100	Desulfurococcales (Crenarcheota)	Yellowstone hot springs
			Staphylothermus achaiicus (AJ012645)	94		Geothermal vents, Greece
PS8SGXN439	3.2	FJ446514	*Thermococcus acidaminovorans* strain DSM 11906 (AY099170)	100	Thermococcales	Hydrothermal deep-sea vents, Italy
			Archaeon enrichment culture clone EA3.5 (EU573147)	99		Ekofisk oil field
PS8SGXN687	3.2	FJ446515	Uncultured archaeon clone SSE_L4_H05 (EU635920)	99	Desulfurococcales (Crenarchaeota)	Nevada hot spring sediment
			Thermosphaera aggregans (X99556)	99		Yellowstone hot spring
PS8SGXN753	3.2	FJ446516	Uncultured archaeon G04b_L4_A09 (EU635911)	99	Thermoproteales (Crenarchaeota)	Nevada hot spring sediment
			Vulcanisaeta distributa strain IC-065 (AB063639)	87		Japan hot spring
PS2SGXN537	0.8	FJ446510	Uncultured crenarchaeote WIP_20m_6B_A (EF420183)	99	Desulfurococcales (Crenarcheota)	Canadian oil sands tailings pond
			"*Desulfurococcus kamchatkensis*" (EU167539)	86		Kamchatka hot spring
PS1SGXN453	0.4	FJ446505	Archaeon enrichment culture clone PW30.6A (EU573155)	99	Archaeoglobales	Ekofisk oil field
			Archaeoglobus profundus (AF297529)	98		
PS1SGXN470	0.4	FJ446506	*Thermococcus mexicalis* strain GY 869 (AY099181)	99	Thermococcales	Hydrothermal deep-sea vents
			Thermococcus sibiricus (AJ238992)	99		Siberian high-temperature oil reservoir
PS1SGXN592	0.4	FJ446507	Archaeon enrichment culture clone PW15.7A (EU573156)	99	Archaeoglobales	Ekofisk oil field
			Archaeoglobus profundus (AF297529)	90		
Coalescer: total # sequences = 240, primers ARC333 and 958R, 97% OTUs from DOTUR						
CO146FGIO3	60.8	FJ446519	Uncultured archaeon SSE_L4_E01(EU635901)	99	Archaeoglobales	Hot spring sediment

			Archaeon enrichment culture clone EA8.8 (EU573151)	98		Ekofisk oil field
			Archaeoglobus fulgidus strain L3 (DQ374392)	98		
CO64FGIO506	26.7	FJ446522	*Thermococcus alcaliphilus* DSM 10322 (AB055121)	100	Thermococales	
			Archaeon enrichment culture clone PW5.2A (EU573152)	99		Ekofisk oil field
CO12FGIO387	5.0	FJ446518	Uncultured archaeon SSE_L4_E01(EU635901)	96	Archaeoglobales	Hot spring sediment
			Ferroglobus placidus (AF220166)	96		
CO10FGIO395	4.2	FJ446517	Archaeon enrichment culture clone PW30.6A (EU573155)	99	Archaeoglobales	Ekofisk oil field
			Archaeoglobus profundus (AF297529)	98		
CO6FGIO519	2.5	FJ446523	Uncultured *Methanothermobacter* sp. clone ARCA-3F (EU073827)	99	Methanobacteriales	Coal enrichment culture
			Methanothermobacter thermautotrophicus strain JZTM (EF100758)	99		Jiaozhou Bay sediment
			Methanothermobacter wolfeii strain KZ24 (DQ657904)	99		Dagang oil field
CO1FGIO425	0.4	FJ446520	Archaeon enrichment culture clone PW30.6A (EU573155)	96	Archaeoglobales	Ekofisk oil field
			Ferroglobus placidus (AF220166)	96		
CO1FGIO557	0.4	FJ446521	Uncultured crenarchaeote Clone MDS-r-E06 (AB353218)	98	Desulfurococcales (Crenarchaeota)	Mesophilic digested sludge
			"*Desulfurococcus kamchatkensis*" strain 1221n (EU167539)	86		Kamchatka hot spring

* most similar sequence and/or isolate in Genbank, as determined by BLASTN.

** Class affiliation, as determined by Classifier (RDP).

One sequence from each operational taxonomic unit (OTU) at the 97% level of similarity as defined by DOTUR *(10)* was chosen from among the sequences in that OTU and its taxonomic affiliation and closest GenBank match was determined. Representative sequences are named with the first two letters indicating the sample origin, the following numerals designate the total number of sequences within that particular OTU. The final four letters and 3-4 numbers are the JGI code identifying the sample library and location within the 384-well plate.

Table S2. Measures of genetic diversity and species richness.

Sample code	# clones	# 99.9% OTUs (%)	# 97% OTUs (%)	S_{Chao1} (97%) (95% CI)	ACE (97%) (95% CI)	H (97%)	S (97%)	% library coverage (97%)
SWBAC	215	91	17	38 (22.1-103.3)	27.2 (19.4-59.8)	1.86 (1.71-2.01)	0.2205	97.2
2PBAC	334	35	9	10 (9.09-19.7)	13.9 (9.8-36.8)	0.47 (0.36-0.59)	0.8139	99.2
2LBAC	648	280	28	54 (34.8-127.2)	45.8 (33.7-83.7)	1.21 (1.08-1.34)	0.5474	98.0
PSBAC	544	267	35	61.3 (42.4-127.5)	53.2 (41.5-86.2)	1.76 (1.63-1.91)	0.3567	97.2
COBAC	329	60	35	65 (43.7-138.2)	56.1 (42.4-95.0)	2.36 (2.21-2.51)	0.1633	95.1
2PARC	226	82	7	7.5 (7.0-15.3)	10 (7.5-25.4)	1.41 (1.34-1.49)	0.2608	99.1
PSARC	252	83	13	14.5 (13.2-28.1)	15.9 (13.4-33.5)	2.00 (1.89-2.10)	0.1684	98.8
COARC	240	63	7	8 (7.07-20.8)	9.2 (7.3-26.5)	1.07 (0.95-1.20)	0.4418	99.2

SWBAC: Bacterial 16S rRNA library from seawater pig envelope sample.

2PBAC: Bacterial 16S rRNA library from well 2P sample.

2LBAC: Bacterial 16S rRNA library from well 2L sample.

PSBAC: Bacterial 16S rRNA library from 1st stage separator sample.

COBAC: Bacterial 16S rRNA library from coalescer sample.

2PARC: Archaeal 16S rRNA library from well 2P sample.

PSARC: Archaeal 16S rRNA library from 1st stage separator sample.

COARC: Archaeal 16S rRNA library from coalescer sample.

OTUs, SChao1 (97%), ACE (97%), H (Shannon-Weaver), S (Simpson's index) were estimated using DOTUR *(10)*.

% library coverage was estimated by the method of Good *(11)* for sequences at 97% similarity.

95% CI: 95% confidence interval values.

Table S3. Core and dominant bacterial OTUs (97% similarity) and physiologies.

OTU closest match (% similarity)	%*	Physiology of closest match	2L	2P	PS	CO	SW	Total #
		Bacteria: Core 2L, PS, CO						
Thermovirga lienii (DQ071273)	99	[17***] Fermentative, H_2S from cystine/S^o	540**	17	313	105	0	975
Desulfomicrobium thermophilum (AY464939)	98	[18]SRB	27	0	39	60	0	126
Thermosipho africanus/geolei (DQ647057/AJ272022)	99	[19]H_2S from S^O	10	0	11	2	0	23
Uncultured clone/*Pelobacter carbinolicus* (AB243821/CP000142)	97	Fermentative, Syntroph, [19]H_2S from cystine/S^O, IR	6	0	20	46	0	72
Desulfacinum subterraneum (AF385080)	97	[20]SRB	7	0	6	6	0	19
Thermodesulfobacterium commune (AF418169)	98	[20]SRB	3	0	6	11	0	20
Thermoanaerobacter pseudethanolicus/X514 (CP000924/CP000923)	99	[21]Fermentative, IR (some strains), thiosulfate reduction	14	2	74	26	0	116
Thermacetogenium phaeum (AB020336)	98	[22]Acetogen, Syntroph, Sulfate/thiosulfate reduction	6	0	9	5	0	20
Uncultured *Thermacetogenium* clone B11_otu13 (DQ097678)	97	Uncultured	1	0	1	1	0	3
Moorella thermoacetica (AY884087)	86	Fermentative	1	0	4	3	0	8
		Bacteria: dominant 2P						
Petrobacter sp. DM-3 (DQ539621)	99	[23]Fermentative NR	0	327	0	0	0	327
		Bacteria: dominant SW						
Pseudomonas stutzeri/putida	98	[24]NR- and/or HC-degrading (some)	1	0	1	0	143	145

*: % similarity

**: number of sequences per sample

***: reference

SRB: sulfate-reducing bacteria

IR: iron reduction

NR: nitrate reduction

Table S4. Dominant archaeal OTUs (97% similarity) and physiologies.

OTU closest match (% similarity)	%*	Physiology of closest match *(25)*	2P	PS	CO	Total #
Archaea: Core 2P, PS, CO						
Archaeoglobus fulgidus DSM 4304/L3 (AE000782/DQ374392)	98	SRA	66**	56	158	280
Thermococcus alcaliphilus/mexicalis/sibiricus (AB055121/AY099181/AJ238992)	99	Fermentative, H_2S from S^O, IR (*T. sibiricus*)	66	78	64	208
Methanothermobacter thermautotrophicus (EF100758)	99	Methanogen H_2-utilizing	64	29	6	99
Uncultured Methanobacteriaceae clone A1m_OTU 3 (DQ097668)	99	Uncultured, from Dang oil field	26	35	0	61
Archaeoglobus profundus (AF297529)	98	SRA	2	1	11	14

*: % similarity

**: number of sequences per sample

SRA: sulfate-reducing archaea

IR: iron reduction

References for Supporting Online Material

1. L. M. Gieg, K. E. Duncan, J. M. Suflita, *Appl. Environ. Microbiol.***74**, 3022-3029 (2008).

2. M. Klein *et al., J. Bacteriol.* 183, 6028–6035 (2001).

3. B. A. Hales *et al., Appl. Environ. Microbiol.* **62**, 668–675 (1996).

4. T. Z. P. DeSantis *et al., Nucleic Acids Res.* **34**,W394-9 (2006a).

5. T. Huber, G. Faulkner, P. Hugenholtz. *Bioinform.* **20,** 2317-2319 (2004).

6. K. E. Ashelford, N. A. Chuzhanova, J. C. Fry, A. J. Jones, A. J. Weightman, *Appl. Environ. Microbiol.* **71**, 7724-7736 (2005).

7. S. F. Altschul *et al*., *Nucleic Acids Res* **25**, 3389-3402 (1997).

8. T. Z. P. DeSantis *et al., Appl. Environ. Microbiol.* **72**, 5069-5072 (2006b).

9. D. J. Lane, in *Nucleic Acid Techniques in Bacterial Systematics,* E. Stackebrandt , M. Goodfellow, Eds. (Wiley, Chichester, U.K., 1991), pp. 115–174.

10. P. D. Schloss, J. Handelsman, *Appl. Environ. Microbiol*. **71**, 1501-1506 (2005).

11. I. J. Good, *Biometrika* **40**, 237–264 (1953).

12. Q. Wang, G. M. Garrity, J. M. Tiedje, J. R. Cole, *Appl. Environ. Microbiol.* **73**, 5261-5267 (2007).

13. R. S. Tanner, R. S. in *Manual of Environmental Microbiology*, C. J. Hurst et al. Eds. (ASM Press, Washington, DC, ed. 3, 2007), pp. 69-78.

14. M. S. Elshahed, L. M. Gieg, M. J. McInerney, J. M. Suflita. *Environ. Sci. Tech.* **35**, 682-689 (2001).

15. L. M. Gieg, J. M. Suflita, in *Petroleum Microbiology*, B. Ollivier, M. Magot, Eds. (ASM Press, Washington, D.C., 2005), pp. 337-356.

16. O. Kniemeyer *et al*., *Nature* **449,** 898-901 (2007).

17. H. Dahle, N. K. Birkeland, *Int J Syst Evol Microbiol*. **56**,1539-1545 (2006).

18. F. Thevenieau, M. L. Fardeau, B. Ollivier, C. Joulian, S. Baena, *Extremophiles* **11**, 295-303 (2007).

19. D. R. Lovley, E. J. Phillips, D. J. Lonergan, P. K. Widman, *Appl. Environ. Microbiol.* **61**, 2132–2138 (1995).

20. N.-K. Birkeland, in *Petroleum Microbiology*, B. Ollivier, M. Magot, Eds. (ASM Press, Washington, D.C., 2005) pp. 35-54.

21. B. Ollivier, J.-L. Cayol, in *Petroleum Microbiology*, B. Ollivier, M. Magot, Eds. (ASM Press, Washington, D.C., 2005) pp. 71-88.

22. S. Hattori, Y. Kamagata, S. Hanada, H. Shoun, *Int J Syst Evol Microbiol*. **50**,1601-1609 (2000).

23. M. B. Salinas *et al., Int J Syst Evol Microbiol*. **54**, 645-649 (2004).

24. N. J. Palleroni, in *Bergey's Manual of Systematic Bacteriology,* D. J. Brenner, N. R. Krieg, J. T. Staley, G. M. Garrity, Eds. (Springer-Verlag. New York, NY, ed. 2, 2005) Vol 2, 323–379.

25. C. Jeanthon, O. Nercessian, E. Corre, A. Grabowski-Lux, in *Petroleum Microbiology*, B. Ollivier, M. Magot, Eds. (ASM Press, Washington, D.C., 2005) pp. 55-70.

The work conducted by the U.S. Department of Energy Joint Genome Institute is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.