RESEARCH ARTICLE

Field and laboratory studies on the bioconversion of coal to methane in the San Juan Basin

Boris Wawrik[1], Margarita Mendivelso[1,2], Victoria A. Parisi[1,2], Joseph M. Suflita[1,2], Irene A. Davidova[1,2], Christopher R. Marks[1], Joy D.Van Nostrand[1,3], Yuting Liang[1,3], Jizhong Zhou[1,3], Brad J. Huizinga[4], Dariusz Strąpoć[5] & Amy V. Callaghan[1,2]

[1]Department of Botany and Microbiology, University of Oklahoma, Norman, OK, USA; [2]Institute for Energy and the Environment, University of Oklahoma, Norman, OK, USA; [3]The Institute for Environmental Genomics, University of Oklahoma, Norman, OK, USA; [4]ConocoPhillips Corporation, Houston, TX, USA; and [5]Dariusz BioGeoChem LLC, Houston, TX, USA

Correspondence: Boris Wawrik and Amy V. Callaghan, Department of Botany and Microbiology, University of Oklahoma, 770 Van Vleet Oval, Norman, OK 73019, USA. Tel.: 405 325 8596; 405 325 1872; fax: 405 325 7619;
e-mail: bwawrik@ou.edu; acallaghan@ou.edu

Received 7 July 2011; revised 18 November 2011; accepted 28 November 2011.
Final version published online 11 January 2012.

DOI: 10.1111/j.1574-6941.2011.01272.x

Editor: Tillmann Lueders

Keywords
coal-bed methane; methanogenesis; metabolite analysis; subsurface microbiology.

Abstract

The bioconversion of coal to methane in the San Juan Basin, New Mexico, was investigated. Production waters were analyzed via enrichment studies, metabolite-profiling, and culture-independent methods. Analysis of 16S rRNA gene sequences indicated the presence of methanogens potentially capable of acetoclastic, hydrogenotrophic, and methylotrophic metabolisms, predominantly belonging to the Methanosarcinales and Methanomicrobiales. Incubations of produced water and coal readily produced methane, but there was no correlation between the thermal maturity and methanogenesis. Coal methanogenesis was greater when samples with a greater richness of Firmicutes were utilized. A greater archaeal diversity was observed in the presence of several aromatic and short-chain fatty acid metabolites. Incubations amended with lactate, hydrogen, formate, and short-chain alcohols produced methane above un-amended controls. Methanogenesis from acetate was not observed. Metabolite profiling showed the widespread occurrence of putative aromatic ring intermediates including benzoate, toluic acids, phthalic acids, and cresols. The detection of saturated and unsaturated alkylsuccinic acids indicated *n*-alkane and cyclic alkane/alkene metabolism. Microarray analysis complemented observations based on hybridization to functional genes related to the anaerobic metabolism of aromatic and aliphatic substrates. These data suggest that coal methanogenesis is unlikely to be limited by methanogen biomass, but rather the activation and degradation of coal constituents.

Introduction

Over the past two decades, coal-bed methane (CBM) has emerged as an important source of fossil energy. While CBM was originally thought to be of thermogenic origin, it is now evident that microbial methanogenesis is significant in many formations (for a review, see Strąpoć *et al.*, 2011). That is, CBM is abundantly found in formations that have never been subject to conditions conducive to thermogenic methane formation. Isotopic analysis of gases from coal formations of varying maturity often indicates mixed signals that suggest both biological and abiotic origins for CBM (Zhou *et al.*, 2005; Strąpoć *et al.*, 2011). Elevated paleo-temperature regimes may have constrained subsurface microbial activity in coal formations at some point in geological history, but the subsequent re-inoculation of shallow areas in North America (Strąpoć *et al.*, 2010), Australia (Faiz & Hendry, 2006), and Asia (Shimizu *et al.*, 2007) has been hypothesized to allow microbial CBM formation to proceed.

These observations have prompted investigations into the ecology of subsurface communities associated with coal formations and of the environmental factors influencing microbial activity. Enrichment studies have mainly focused on the detection of methanogenic archaea in formation water. For example, in studies of coal-mine water collected from the Ruhr basin, it was shown that methane was produced when samples were incubated in the

FEMS MICROBIOLOGY ECOLOGY

presence of H_2 or fatty acids (Thielemann *et al.*, 2004). Enrichments using hard Ruhr basin coal and mine timber produced methane over a 9-month period, and acetoclastic methanogenesis was observed in enrichments containing acetate (Krüger *et al.*, 2008). Similar observations were made for samples collected from the Powder River Basin, which yielded active methane production in the presence of H_2/CO_2 (Harris *et al.*, 2008). As coal is hypothesized to be the ultimate source of CBM, enrichment studies using coal as a sole carbon substrate have also been conducted and have demonstrated that ambient microbial communities can actively produce methane from coal (Green *et al.*, 2008; Harris *et al.*, 2008). While these studies collectively indicate that the requisite methanogens are present in coal seams, the physiological range of coal-bed methanogens, particularly in the context of other participating organisms, and the ecological conditions that ultimately lead to methanogenesis still remain unclear.

Molecular approaches have provided some insight into the phylogenetic composition of methanogenic archaeal communities in coal via the analysis of PCR-amplified 16S rRNA gene sequences (Shimizu *et al.*, 2007; Klein *et al.*, 2008; Li *et al.*, 2008; Strąpoć *et al.*, 2008; Fry *et al.*, 2009; Penner *et al.*, 2010), but these investigations included small sequence data sets, typically obtained from only a single sample. A comprehensive survey of the 'core taxa' found across the spatial context of a methane-producing coal bed has not yet been reported, and the present effort seeks to provide some of this information.

Coal is composed of a complex mixture of aromatic, heterocyclic, and aliphatic constituents and is thought to be broken down by a sequence of primary and secondary fermenters, providing short-chain organic acids (e.g. acetate), alcohols, and H_2 for acetoclastic, methylotrophic, and hydrogenotrophic methanogenesis, respectively (Strąpoć *et al.*, 2011). The mechanisms that microorganisms utilize to initiate the degradation of coal organic matter remain unclear, but a growing body of literature indicates that aliphatic and aromatic coal 'monomers' can be degraded under anoxic conditions in subsurface environments. The salient mechanisms are still emerging, but several common pathways appear to be prevalent, including addition to fumarate, carboxylation, hydroxylation, and methylation (Abu Laban *et al.*, 2009; Boll & Heider, 2010; Tierney & Young, 2010; Widdel & Grundmann, 2010; Widdel & Musat, 2010). In many cases, these reactions produce characteristic metabolites that are not known to form abiotically (e.g. Parisi *et al.*, 2009). The detection of diagnostic metabolites can provide unequivocal evidence for *in situ* microbial activity. In some cases, the requisite enzymes (and the genes that encode them) are known and can serve as similar biomarkers for the genetic potential of microbial communities to utilize specific substrates. For example, the addition of aromatic and aliphatic hydrocarbons to fumarate is catalyzed by several classes of glycyl radical enzymes, such as benzylsuccinate synthase (Bss) (for review, see Widdel & Musat, 2010), (2-naphthylmethyl)succinate synthase (Nms) (Musat *et al.*, 2009) and alkylsuccinate synthase (Ass) [also known as (1-methylalkyl)succinate synthase (Mas)] (Callaghan *et al.*, 2008; Grundmann *et al.*, 2008), respectively. The dehydrogenation (anaerobic hydroxylation) of ethylbenzene is catalyzed by ethylbenzene dehydrogenase (Ebd) (Johnson *et al.*, 2001; Kniemeyer & Heider, 2001), and more recent studies have proposed the carboxylation of benzene and naphthalene via the putative enzymes, benzene carboxylase (Abc) (Abu Laban *et al.*, 2010) and naphthalene carboxylase (Bergmann *et al.*, 2011), respectively. Combining metabolic profiling with the detection of diagnostic genes (Beller *et al.*, 2008; Callaghan *et al.*, 2010; Yagi *et al.*, 2010) can provide a powerful screen for assessing the underlying biochemical mechanisms of coal biodegradation. A detailed survey of a coal environment that combines these approaches is reported here.

The key challenges to understanding CBM formation are elucidating the mechanisms of microbial coal activation, identifying the requisite microorganisms, and determining the physiological constraints on the system (Strąpoć *et al.*, 2011). Prior studies have focused on individual aspects of these questions. In this study, an integration of several approaches was taken to investigate microbial CBM formation in the San Juan Basin (SJB). We conclude that the SJB contains a wide diversity of methanogens and that the anaerobic microbial communities in the SJB exhibit an impressive ability to transform coal organic matter, as well as a series of model substrates. Therefore, CBM is not likely to be restricted by methanogen biomass but by the slow anaerobic bioconversion of recalcitrant coal organic matter constituents.

Materials and Methods

Sampling sites

Sampling was conducted at eleven well sites located along the Colorado-New Mexico border, within the Upper Cretaceous Fruitland Coal Formation (Fig. 1). Production water samples for cell counts and DNA extraction were obtained directly from well heads and processed in the field (see below). Inocula for enrichment studies and metabolite-profiling samples were collected in degassed bottles (see below), stored and shipped to the laboratory in coolers containing ice packs, and then stored at 4 °C until use.

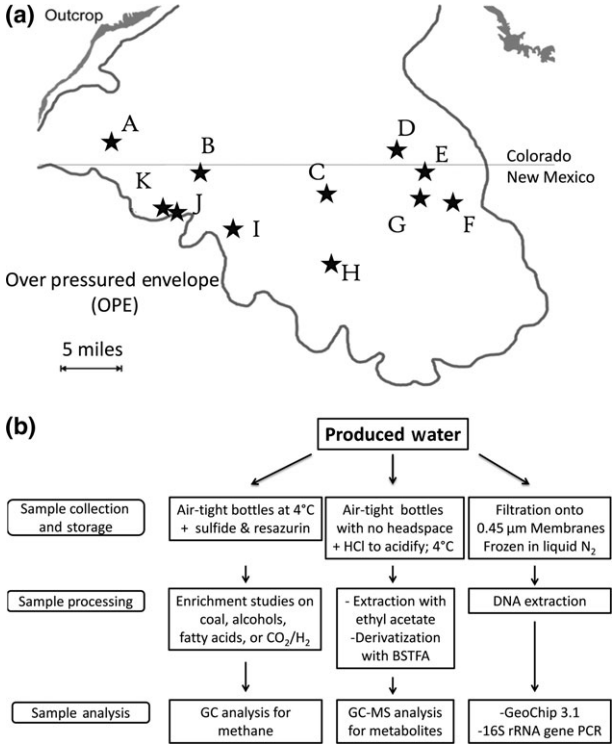

Fig. 1. (a) Locations of wells that were sampled relative to the over-pressured envelope within the Fruitland Coal formation of the SJB. The horizontal line indicates the border between Colorado (CO) and New Mexico (NM). Grey shaded areas indicate outcrops of Fruitland Coals, which were sampled to obtain coal sample 'A' (Table 2). (b) Schematic diagram of sampling and analysis performed in this study.

Sample collection and incubations

Produced water samples were collected into 1-L degassed (N_2), sterile Schott bottles containing 1 mL of resazurin (1 g L^{-1}) and 10 mL of sodium sulfide (0.5 mM). The bottles were stored at 4 °C until they were processed inside an anaerobic glove box. Microbial incubations were established under anaerobic conditions in a bicarbonate-buffered reduced minimal medium (Parisi *et al.*, 2009). The samples of produced water were concentrated 10 times by centrifugation under anaerobic conditions. The concentrated suspensions were used as inocula. Composite inocula were generated based on the chemical composition of produced waters. Composite inoculum I was generated from wells A, D, E, F, and I, which contained low levels of chloride and sulfate (0–17 mM Cl^-, < 1 mM SO_4^{2-}). Composite II was generated from wells B, C, H, and J, which exhibited medium levels of chloride and low levels of nitrate (30–50 mM Cl^-, < 1 mM NO_3^-), while composite III was generated from the remaining water samples G and K, which displayed high levels of chloride and pH (80–140 mM Cl^-, pH 7.5–9).

Four different Fruitland Coal samples (labeled A–D) of varying thermal maturities were collected. Coal A (0.6% R_o) originated from an outcrop of Fruitland Coal near Durango, CO. The remaining coal samples were retrieved from cores collected by ConocoPhillips as part of drilling operations in the SJB. Coals B and C (0.87% R_o) had similar thermal maturities and were therefore combined. Coal D had the highest thermal maturity among the collected material (1.1% R_o). Upon arrival in the laboratory, the coals were exposed to air for several hours during mechanical processing using mortar and pestle. The amount of coal D available was insufficient to construct assays with each of the eleven well waters. Therefore, it was only possible to compare the methanogenic activities among the composite inocula that all received the limited amount of coal D.

Incubations were established in an anaerobic glove box under an atmosphere of N_2. Sterile serum bottles (160 mL) received 2 g of sterile (autoclaved), ground coal, 2 mL of concentrated inoculum (either from individual wells or composite samples) and 20 mL of reduced, modified mineral salts medium (pH 7.1–7.3) (McInerney *et al.*, 1979). The bottles were closed with butyl rubber stoppers, secured with aluminum crimp seals and incubated under a N_2 : CO_2 (80 : 20) headspace at 31 °C. Strict anaerobic technique was used for all culture manipulations as well as for media and substrate preparations. Substrate un-amended (e.g. coal free) controls were included to account for background levels of methane production. Un-inoculated, sterile (autoclaved at 121 °C for 20 min, three times), lactate-amended (25 mM), and toluene-amended (0.4 mM) incubations were also included as experimental controls.

Enrichments on short-chain fatty acids, alcohols, and hydrogen were conducted in minimal medium containing (per L): 20 mL of B Salt Mix (2.14 M NaCl, 40 mM $MgSO_4$, 34 mM KCl, 23 mM CaCl), 100 mL of BT Salt Mix (295 mM Na_2HPO_4, 110 mM KH_2PO_4, 56 mM NH_4Cl) (Taylor *et al.*, 1970), 1 mL of 1 M $NaHCO_3$, 1 mL of selenite–tungstate solution (10 mM NaOH, 11.6 µM Na_2SeO_3, 12.1 µM Na_2WO_4), 1 mL of SL-10A trace metal solution (Widdel & Bak, 1992), and 1 mL of resazurin (1 g L^{-1}). The medium was degassed under a stream of N_2 : CO_2 (70% : 30%) for 1 h. Forty milliliters were dispensed into 55-mL serum bottles, which were sealed using butyl rubber stoppers. Each bottle was amended with 0.5 mL of anaerobic cysteine sulfide solution (12.5 g L^{-1} NaOH, 25 g L^{-1} cysteine·HCl, 25 g L^{-1} N_2S). Bottles were sterilized by autoclaving, and 0.1 mL of sterile-filtered, anaerobic VB-1 vitamin solution (Widdel & Bak, 1992) (modified to

include 50 mg of vitamin B_{12}) was added. Ethanol, 1-propanol, 1-butanol, formate, acetate, and propionate (10 mM) were amended from sterile, anaerobic stocks. For enrichments incubated in the presence of hydrogen, the headspace was flushed with H_2 for 1 min and then pressurized with H_2 to 20 psi. Each bottle received two milliliters of respective formation water as an inoculum. The pH of each bottle was checked to make sure that it was not altered by substrate and inoculum addition. Heat-killed controls were prepared for each well site and for each carbon source by autoclaving a duplicate set of bottles. Substrate-free controls (NCC) were also generated for each of the sites.

Cell counts

Triplicate samples (10 mL) of produced water were preserved on-site with 3.7% filtered formaldehyde in sterile, disposable 15-mL tubes and stored at 4 °C until analysis. Cells were stained with DAPI (4′-6-diamidino-2-phenylindole) (final concentration of 5 μg mL^{-1}). Samples were incubated in the dark for 30 min at room temperature and then filtered onto 0.2-μm black polycarbonate membranes (Sterlitech, Kent, WA). Cells were then enumerated via epifluorescence microscopy.

DNA extraction

Produced water was filtered immediately in the field onto 25-mm 0.45-μm Supor filter membranes (Pall Life Sciences, Ann Arbor, MI). Owing to the varying nature of the produced fluids (i.e. particulates, the presence of gas condensate, etc.), filtration was conducted until the filter capacity was achieved. The filtration volume for each sample was recorded. Filters were stored in 1× STE buffer (1 M NaCl, 100 mM Tris–HCl pH 8.0, 10 mM EDTA pH 8.0), immediately frozen in liquid N_2 in the field, and stored at −80 °C until extraction. DNA extraction was performed as previously described (Rainey *et al.*, 1996; Callaghan *et al.*, 2010).

PCR

PCR for bacterial and archaeal 16S rRNA genes was conducted using primers spanning the V1 and V2 variable regions of the 16S rRNA gene. Bacterial primers were 27F (5′-AGAGTTTGATCCTGGCTCAG-3′) and 338R (5′-TGCTGCCTCCCGTAGGAGT-3′). The archaeal primers were A8F (5′-TCCGGTTGATCCTGCC-3′) and A344R (5′-TCGCGCCTGCTGCICCCCGT-3′). Reactions contained 1 μL of DNA template (5–10 ng), 0.5 μL of each primer (100 nM of each primer), and 45 μL of SuperMix (Invitrogen, Carlsbad, CA). PCR conditions were as follows: 95 °C for 2 min, followed by 35 cycles of 96 °C for 30 s, 55 °C for 30 s, and 72 °C for 60 s. PCRs were purified using a QIAquick PCR purification column (Qiagen, Valencia, CA) to remove primers. Purified PCR products were visualized by gel electrophoresis to confirm amplification and to verify negative controls. From each of the purified PCRs, 5 μL was added to a second PCR containing barcoded PCR primers for multiplexed pyrosequencing as described by Hamady *et al.* (2008). Each sample received a different tagged forward primer, containing a specific 'barcode' sequence, and samples were 'tagged' by re-amplification for six cycles. Tagging was confirmed by comparing the sizes of tagged and un-tagged PCR products by gel electrophoresis. PCR products were mixed and sequenced via parallel 454-sequencing at the University of South Carolina's sequencing core facility (http://engencore.sc.edu/). Computational analysis was used to sort the sequences based on their tags to assign them to specific well sites. Sequence analysis of high-quality pyrosequencing reads was performed using MOTHUR (Schloss *et al.*, 2009). All sequences were aligned to the SILVA Reference Alignment using the NAST algorithm (Pruesse *et al.*, 2007). Only sequences lacking ambiguous bases, containing minimal homopolymers ($n > 8$), having an average quality of at least 25, and aligning to the appropriate gene regions were retained for analysis. All sequences were checked for possible chimerism using an internal algorithm based on ChimeraSlayer (www.broadinstitute.org), and all putative chimeras were removed. Retained sequences were preclustered using a single-linkage clustering algorithm prior to pairwise distance calculations to reduce pyrosequencing error. Operational taxonomic units (OTUs) were generated at 10% sequence identity. A 10% cutoff was viewed as a conservative choice. The PCR products that were used in pyrosequencing were relatively short (340 bp), and species-level classification based on short reads can be weak. This is especially true for the archaea where taxonomy is less robust. A 10% cutoff corresponds to a family-level classification, representing a compromise, balancing the need for robust classification with community-level resolution. A single sequence from each OTU was selected at random for phylogenetic and taxonomic representation of the corresponding OTU. Representative sequences were classified using the Ribosomal RNA Database Project's naive Bayesian rRNA classifier (Wang *et al.*, 2007) and verified phylogenetically by parsimonious methods with the SILVA SSU Reference Database (Pruesse *et al.*, 2007) in ARB (Ludwig *et al.*, 2004).

Accession numbers

Representative sequences for each of the OTUs obtained by 454 pyrosequencing were deposited into the NCBI

GenBank archive with the following accession numbers: JF274498–JF2774598.

Metabolite profiling

Produced water (1 L) was collected into sterilized 1-L bottles and immediately acidified with 50% HCl (pH < 2) to protonate acid intermediates and halt microbial activity. Samples were transported back to the laboratory and stored at 4 °C until extraction with ethyl acetate. Extracts were dried over anhydrous Na_2SO_4 and concentrated to a volume of 100 μL (10 000 fold concentration) by rotary evaporation and N_2. All concentrated extracts were derivatized with *N,O*-bis(trimethylsilyl)trifluoroacetamide (Sigma Aldrich, St. Louis, MO) prior to gas chromatography–mass spectrometry (GC–MS) analysis. The derivatized organic extracts were analyzed with an Agilent 6890 model GC coupled with a model 5973 MS and separated with a DB-5ms capillary column (30 m × 0.25 mm i.d. × 1.0 μm film; J&W Scientific, Folsom, CA). The oven temperature was held at 45 °C for 5 min and increased at 4 °C/min to 270 °C and held for 10 min before mass spectral analysis (Duncan *et al.*, 2009). All metabolite identifications were made by comparison with authentic derivatized standards (purchased commercially or synthesized) or with previously reported MS profiles of over 90 compounds implicated in anaerobic hydrocarbon metabolism (Elshahed *et al.*, 2001; Kniemeyer *et al.*, 2007; Duncan *et al.*, 2009; Gieg *et al.*, 2009). The chromatographic peaks of positively identified metabolites were analyzed with respect to their integrated area using the MS ChemStation (G1701DA D.01.00) Software (Agilent Technologies, Santa Clara, CA). For most compounds, the limit of detection for the GC–MS analysis was approximately 20 μM (Duncan *et al.*, 2009). The presence or absence of a metabolite, as well as the relationships among samples, was determined using CLUSTER 3.0 (http://bonsai.ims.utokyo.ac.jp/~mdehoon/software/cluster/software.htm) (de Hoon *et al.*, 2004) and JAVA TREE VIEW (http://jtreeview.sourceforge.net) (Saldanha, 2004) open source software packages for dendrogram and heat map creation. Euclidean distances and average linkages were used in the hierarchical clustering analysis to describe the relationship among samples.

Methane measurements

All methane measurements were conducted by injecting 0.2 mL of the enrichment bottle headspace into a Varian 3300 gas chromatograph (GC) equipped with a flame ionization detector and packed stainless steel column (Poropak Q, 80/100; Supelco, Bellefonte, PA). The injector and column temperatures were held at 100 °C, and the detector was 125 °C.

GeoChip analysis

The GeoChip 3.1 is a functional gene array that contains approximately 28 000 probes for more than 56 500 gene targets in 292 gene families (He *et al.*, 2010). DNA hybridization on the GeoChip was performed as previously described (Wu *et al.*, 2006). In brief, whole microbial community DNA (100 ng) was amplified using a Templiphi kit (GE Healthcare, Piscataway, NJ) by applying a modified reaction buffer containing 0.1 μM spermidine and 260 ng μL^{-1} single-stranded binding protein. This improves amplification efficiency and representativeness (Wu *et al.*, 2006). The DNA (~ 2.5 μg amplification product) was then labeled using fluorescent dye, Cy5 (GE Healthcare), by random priming (Wu *et al.*, 2006; Van Nostrand *et al.*, 2009). The DNA was then purified via a QIAquick purification column (Qiagen, Valencia, CA) and resuspended in 50 μL of hybridization solution containing 40% formamide, 3× SSC (450 mM sodium chloride, 45 mM sodium citrate, at pH 7.0), 10 μg of unlabeled herring sperm DNA (Promega, Madison, WI), and 0.1% SDS. Denaturation at 95 °C for 3 min was followed by cooling to 65 °C, and samples were then added onto a GeoChip array. Hybridizations were conducted with a MAUI Hybridization System (BioMicro systems, UT) according to the manufacturer's recommendations (at 42 °C for 10 h). Microarrays were scanned using a ScanArray Express Microarray Scanner (Perkin Elmer, Boston, MA) at 633 nm. The IMAGENE v. 6.1 (Biodiscovery, El Segundo, CA) software was used to determine spot intensity. A signal-to-noise ratio [(signal intensity background)/standard deviation of background)] of < 2.0 was applied to remove poor-quality spots (He & Zhou, 2008). A universal Cy3 standard in each subgrid was used to normalize the Cy5 intensity in respective subgrids.

Results

The SJB is the most prolific methane-producing coal bed in North America, generating more than 26.2 billion cubic meters of natural gas per year (EPA, 2004). Methane originates from the coals of the Upper Cretaceous Fruitland Formation, and most production occurs in the northern area of the basin, where active groundwater recharge leads to hydrostatic over-pressurization. All samples analyzed in this study were collected from wells within this over-pressured envelope (Fig. 1a). The produced water contained low levels of sulfate and nitrate (avg: 200 and 140 μM, respectively), indicating limited

Table 1. Methane production (μM CH_4) in enrichments with methanogenesis substrates and coal

(a)

Well	Incubation time (days)	Methanol	Ethanol	Propanol	Butanol	Formate	Acetate	Propionate	H2	HKC	NCC
A	131	8.0	36.9	23.6	0.1	15.3	0.5	0.0	79.7	–	0.1
B	131	–	–	–	–	–	–	–	–	–	–
C	131	0.1	–	0.1	155	1.0	0.1	0.1	–	0.1	3.7
D	131	–	104	75.3	1.6	0.1	–	0.0	4.8	–	–
E	131	209	66.0	17.5	69.6	14.8	0.3	0.4	17.6	–	0.3
F	131	2.2	1.3	12.1	393	2.7	0.6	–	0.8	–	0.2
G	131	0.1	55.7	7.9	0.1	0.7	0.1	0.1	2.5	0.1	0.1
H	131	237.0	48.3	68.8	0.4	38.2	0.2	0.1	129	–	1.7
I	131	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
J	131	539	32.5	125	20.4	226	0.8	0.1	16.0	–	1.6
K	131	0.1	0.3	0.0	0.1	0.1	0.1	0.1	1.1	–	0.1

(b)

Well	Incubation time (days)	Lactate	Toluene	Heizer/Vanderslice coal	Fruitland coal	Reese Mesa coal
A	288	63.2	–	–	3.2 ± 6	n.a.
B	268	–	–	–	–	n.a.
C	275	124	–	–	6.3 ± 5	n.a.
D	282	11	–	–	–	n.a.
E	275	46.1	–	–	–	n.a.
F	282	10.3	–	–	–	n.a.
G	260	13.2	–	–	3.4 ± 6	n.a.
H	268	88	–	3.4 ± 6	19.8 ± 8	n.a.
I	288	129	10	–	14.2 ± 2	n.a.
J	268	112	11.3	6.8 ± 6	11.9 ± 1	n.a.
K	260	11.1	–	–	–	n.a.
Composite I*	343	110	11	n.a.	n.a.	8.7 ± 8
Composite II†	243	119	10.1	n.a.	n.a.	18 ± 4
Composite III‡	243	13	–	n.a.	n.a.	179 ± 296

HKC, heat-killed controls; NCC, no carbon source amendment control; '–', no methane detected.
*0–17 mM Cl^-, < 1 mM SO_4^{2-}.
†30–50 mM Cl^-, < 1 mM NO_3^-.
‡80–140 mM Cl^-, pH 7.5–9.

potential for sulfate and nitrate respiration. Chlorinity ranged between 0.3 and 133 mM (fresh water to brackish), and the pH was circumneutral to slightly alkaline, ranging between 7.5 and 9 and averaging 7.7 (Supporting Information, Table S1).

Produced water samples were characterized in multiple ways (Fig. 1b). First, enrichments were established to determine whether ambient communities were capable of producing methane from coals of varying thermal maturities. It was assumed that coal would be metabolized into simpler organic molecules that would ultimately be readily converted to methane. Therefore, inocula from across the SJB were also used in a methanogenic bioassay with small molecular weight fatty acids, alcohols, and hydrogen to test this hypothesis. Secondly, we assayed for the presence of polar organic compounds that might be putative coal metabolites indicative of the types of compounds undergoing active biodegradation. Lastly, a molecular analysis was used to assess the genetic potential for coal biodegradation and to describe the core archaeal and bacterial taxa found across the formation.

Enrichments and methane analysis

Enrichment studies indicated that six of eleven well samples contained communities capable of methanogenesis when incubated in the presence of coals of intermediate thermal maturity (R_o = 0.87%, coal samples B/C) (Table 1b). These enrichment cultures produced methane at rates of 0.005–0.04 μM g^{-1} day^{-1} (all sterile and un-inoculated controls showed negligible methane formation;

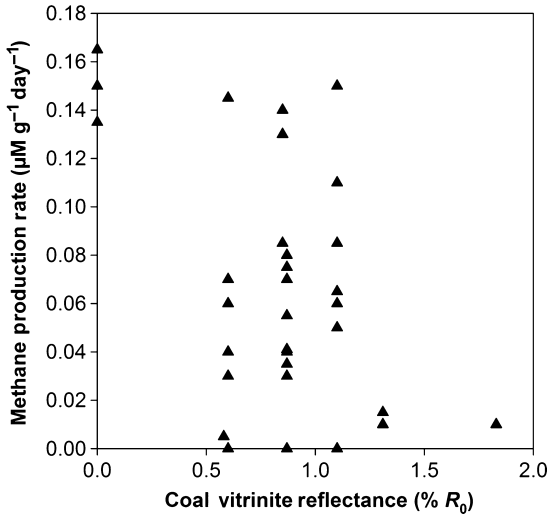

Fig. 2. The rate of methanogenesis in enrichment cultures compared to coal maturity. The rate of methanogenesis was calculated for each enrichment bottle via linear regression of headspace methane concentrations as determined by GC. Data shown here include rates generated from enrichments of SJB coals (Table 1), as well as several additional rates generated from coals of differing maturities. These additional coals were obtained from the United States Geological Survey (USGS) Core repository in Norman, Oklahoma, and incubated in the presence of communities sampled from SJB produced waters. The methanogenesis rate (μM g^{-1} day^{-1}) is plotted against vitrinite reflectance (% R_o).

data not shown). Only two well samples (H and J) produced significant amounts of methane in the presence of coal with low thermal maturity (R_o = 0.6%; coal sample A), but all three composite inocula were capable of producing methane when incubated in the presence of coal with high thermal maturity (R_o = 1.1%; coal sample D). Composite inoculum III was by far the most prolific, producing methane at a rate of 0.37 μM g^{-1} day^{-1}. Overall, these data indicate that methanogenic consortia capable of utilizing coal are widely distributed across the SJB. However, methanogenesis and coal thermal maturity were not highly correlated (Fig. 2).

All samples, except for those from well 'B', produced methane when amended with lactate, H_2, short-chain fatty acids, or alcohols (Table 1a). The greatest average amount of methane was produced in enrichments incubated with methanol, followed by 1-butanol, ethanol, and 1-propanol, respectively. Neither acetate nor propionate served as significant sources of carbon for methanogenesis. On average, enrichments incubated with methanol produced 244- and 1130-fold more methane than those incubated with acetate or propionate, respectively, indicating a clear preference of SJB communities for short-chain alcohols over fatty acids as methanogenesis substrates. Hydrogen and formate were also readily utilized in methanogenic enrichments. These data highlight the importance of hydrogenotrophic and methylotrophic methanogens.

Methanogenic community profiling

Archaeal community composition was profiled using hybridization of community DNA to *mcrA* probes on the GeoChip and by analysis of 16SrRNA gene sequences obtained by 454-based pyrosequencing of PCR products. Twenty-nine positive *mcrA* signals were observed across all samples, 21 of which belonged to uncultivated lineages. The remaining eight positive *mcrA* probes are based on the sequences of cultivated archaea within the Methanomicrobiales and Methanobacteriales (Fig. 3). Among these, hybridization to the probe derived from *Methanocorpusculum labreanun* (GenBank accession# 124363917) was particularly prominent, being present in every sample that was tested. The DNA derived from three samples (wells C, D, and G) accounted for > 72% of all positive *mcrA* probes, suggesting that coals from this area may have harbored the most diverse methanogenic communities. This, however, was not confirmed when 16S pyrosequencing data were examined via rarefaction analysis (Fig. S1).

Archaeal 16S rRNA gene amplicons were obtained from all tested samples except from well 'B'. This is consistent with the observation that formation water from well 'B' did not yield methane in enrichment experiments (see above). A total of 87 515 archaeal 16S reads were generated, 83.5% of which classified within four methanogenic lineages: Methanococcales, Methanosarcinales, Methanomicrobiales, and the uncultured and unclassified RC-I cluster (Erkel *et al.*, 2006) (Fig. 4). To investigate the salient archaeal taxa across the SJB, this analysis was performed for only the 'core taxa' (i.e. OTUs that, at the 10% identity level, comprised at least 1% of any individual 16S sequence library). Thirty-seven different archaeal core taxa were detected (Fig. 4). A clustering analysis was performed (Fig. S2) to investigate which of these taxa were characteristic of the formation. More than 45% of all sequence reads classified within ten OTUs that belonged to the Methanosarcinales. Among these, OTUs 25 and 190 accounted for 28.5% of all data. These two OTUs were very closely related to *Methanosaeta thermophila* (Fig. 4) and were found throughout the formation. OTU 92 is closely related to *Methanosarcina bakeri*, represented 14.2% of all sequence reads, and was also found at all sites where archaeal PCR products were obtained. The Methanomicrobiales accounted for 50% of all detected OTUs, but only 27.5% of all sequence reads, highlighting the relatively

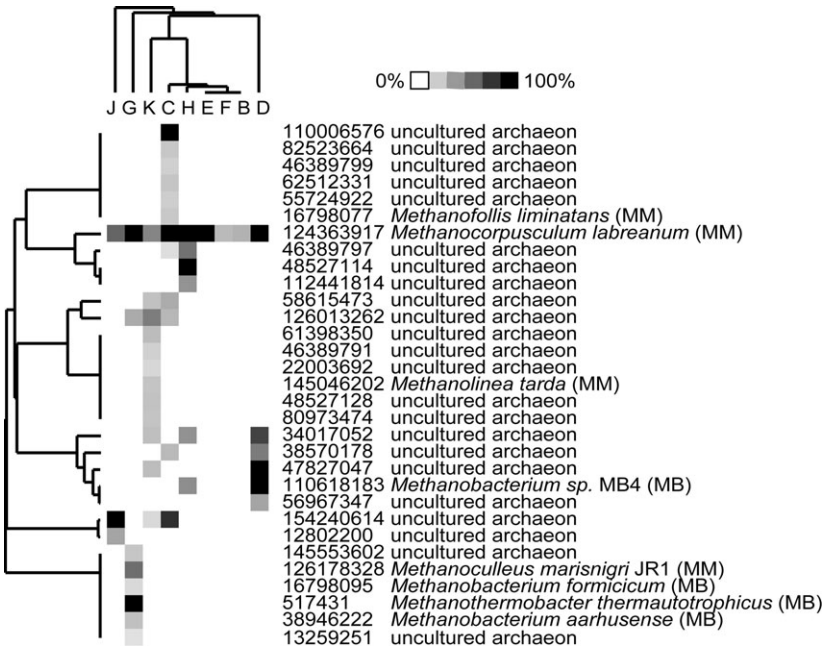

Fig. 3. Analysis of *mcrA* gene probes that produced positive hybridization signals during the functional gene microarray (GeoChip) analysis. The gray scale of the heat map indicates hybridization signal intensity (linear scale), where darker squares indicate greater intensity and are scaled to the highest intensity observed. Letters at the top correspond to the site locations (Fig. 1a). The nearest neighbor dendrogram at the top was generated by calculating Euclidean distances of hybridization signal intensity of each sample. The nearest neighbor dendrogram on the left was calculated based on signal intensity of hybridization across samples. GenBank accession numbers and their corresponding taxonomic annotation are shown to the right of the heatmap.

greater importance of the Methanosarcinales across the formation.

Bacterial community profiling

A comparable analysis was conducted using bacterial 16S rRNA gene PCR products. A total of 56 839 bacterial 16S sequences were assembled into 65 OTUs using the same definition of 'core taxa' as for archaeal libraries. Bacterial core taxa mainly consisted of Actinomycete, Firmicute, and Proteobacterial lineages (Fig. 5). The largest proportion of core taxa (25 OTUs) was found in the Firmicute lineage, among which *Thermoanaerobacterales* was the dominant lineage. Twenty-one OTUs classified within a diverse group of proteobacterial lineages. Some of these classified as known sulfate-reducing and syntrophic bacteria, including Desulfovibrionales, Desulfomonadales, and Syntrophobacterales. Twelve OTUs classified within the Actinomycetales.

Metabolite profiling

Metabolite profiling was used to detect the presence of putative biomarkers indicative of coal biodegradation (Fig. 6). Coal matrices consist of complex mixtures of aromatic and aliphatic hydrocarbons, as well as NSOs (nitrogen-, sulfur- and oxygen-containing heterocyclic compounds) (for review, see Strąpoć *et al.*, 2011). Although the specific pathways of coal activation and subsequent degradation are unclear, methanogenic degradation is thought to proceed via primary fermentation of large molecular weight coal constituents and smaller monomers to intermediates (e.g. fatty acids, organic acids, alcohols, hydrogen and carbon dioxide) that serve as substrates for syntrophs, homoacetogenic bacteria, and methanogens. In this context, the presence/absence of 93 putative metabolites consistent with the microbial metabolism of coal components such as mono-, polyaromatic (PAH), straight chain and alicyclic hydrocarbons was evaluated. The GC–MS analysis of extracts of produced waters resulted in the detection of a series of key polar compounds (Fig. 6). The low molecular weight (C_2–C_9) fatty acids were widely distributed in the SJB formation waters suggesting the microbial metabolism of larger molecular weight constituents. Succinate and adipate were also detected in multiple wells. The larger molecular weight fatty acids (C_{10}–C_{14}) exhibited a far more restricted distribution, whereas hepta- and octadecanoic acids, typical constituents of microbial lipids, were widely present. Although not ubiquitous, the detection of

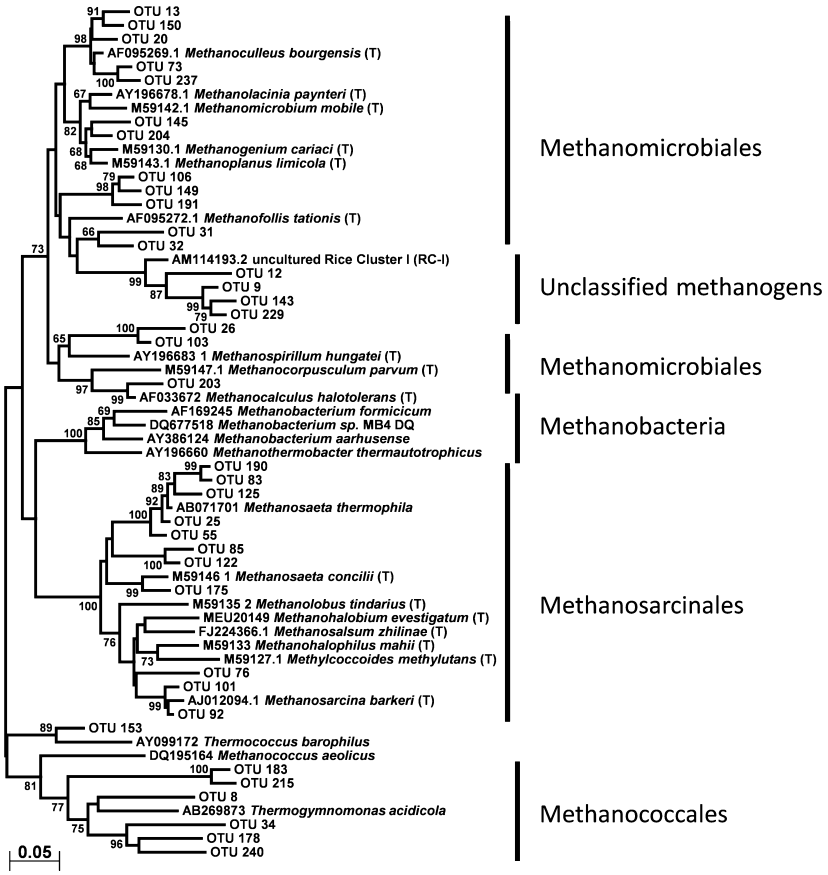

Fig. 4. Phylogenetic analysis of archaeal 16S rRNA gene sequences detected in SJB produced waters. Archaeal 16S PCR products were obtained for each site and sequenced via 454-based pyrosequencing. OTUs were generated at the 10% identity level, and only core taxa are shown (i.e. those OTUs which accounted for at least 1% of sequences in at least one of the sequence libraries). A representative sequence from each OTU was chosen at random using MOTHUR and its closest relative determined using the Ribosomal RNA Database Project (http://rdp.cme.msu.edu/). All type strains for archaeal families within the Methanosarcinales and Methanomicrobiales are shown and designated with a 'T'. Where type strains are not closely related to unknowns, close sequence matches from the database are shown. Sequences were aligned using CLUSTALW, and alignments were manually curated. The phylogeny was constructed using the neighbor-joining approach after calculating distances via the Tajima –Nei method.

saturated and unsaturated alkylsuccinic acids provided evidence of *n*-alkane and cyclic alkane/alkene metabolism (via addition to fumarate) (Gieg & Suflita, 2002; Parisi *et al.*, 2009). The detection of *n*-saturated and methyl-branched fatty acids can also be indicative of alkane biodegradation and/or fatty acid metabolism. Although key intermediates of anaerobic PAH biodegradation were not widespread, several indications of naphthalene metabolism were noted in multiple wells. For example, naphthoic acid and methylated derivatives were found in wells K, A, and D. Additional evidence for the anaerobic biodegradation of naphthalene was supported based upon the presence of the partial ring reduction compound, tetrahydronaphthoic acid in wells G, K, A, and D.

Benzoate was also detected in all samples. Benzoate is a key intermediate in aerobic and anaerobic aromatic compound metabolism (Carmona *et al.*, 2009; Pérez-Pantoja *et al.*, 2010). In addition to benzoate, toluic acids (*o*-, *m*- and *p*-), phthalic acids (*o*-, *m*- and *p*-), and cresols (*m*- and *p*-) were found to be widespread at the site, which suggests that the biodegradation of aromatic-core constituents of the coal matrix is likely an important process in the SJB.

Functional gene analysis

The presence of genetic markers of anaerobic metabolism and the activation of key substrates were investigated using microarray technology (Table 2). A total of 8054 positive probes for functional genes were detected (data not shown) using the GeoChip 3.1. Prominent signals among the relevant genes were those involved in methanogenesis (*mcrA*, and CODH) and sulfate reduction (*aprA*, *dsrA*, and *dsrB*), indicating the presence of a

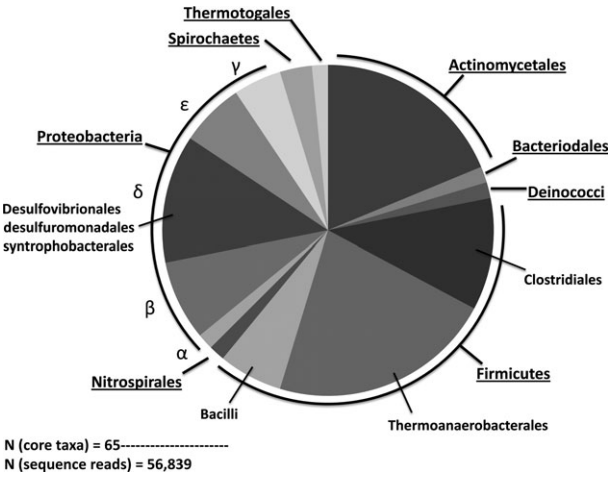

Fig. 5. Phylogenetic breakdown of bacterial 16S rRNA 'core taxa' detected in SJB produced water samples. Bacterial 16S PCR products were obtained for each site and sequenced via 454-based pyrosequencing. Sequences were analyzed using MOTHUR (Schloss *et al.*, 2009) as described in the methods section included in the main text. OTUs were defined at the 10% identity level. All OTUs that accounted for at least 1% of the sequence reads in at least one of the libraries were defined as 'core taxa' and considered for further analysis. Representative sequences from each OTU were chosen at random using MOTHUR, and their phylogenetic affiliations were determined using the GreenGenes database (http://greengenes.lbl.gov/). Underlined taxa indicate bacterial phyla into which core taxa were classified. Nonunderlined taxa are family-level assignments (not all are shown). A class-level phylogenetic breakdown of bacterial 16S rRNA gene sequences can be found in the supporting information (Figure S3).

diverse community of methanogens and sulfate reducers. Six sites displayed hybridization for *assA*, the gene encoding the catalytic subunit of alkylsuccinate synthase, which has been used as a biomarker for anaerobic alkane activation (Callaghan *et al.*, 2010). Also detected were signals for genes associated with benzylsuccinate synthase (*tutD* and *bbsG*) and ethylbenzene dehydrogenase (*ebdBC*), which have been implicated in the activation and subsequent degradation of aromatic substrates such as toluene, xylene isomers, and ethylbenzene. Benzoate is a key intermediate via which many aromatic hydrocarbons are degraded anaerobically. In consistent fashion, there was widespread occurrence of genes coding for the enzymes involved in the anaerobic breakdown of benzoate (*bclA*, *bcr*, *badK*, and *badH*).

Meta-analysis of community data

An attempt was made to correlate community data with the varying parameters that were measured. These analyses included cross correlation analysis, CCA, PCoA, and rarefaction analysis. These attempts did not reveal many interesting patterns (Figure S5), suggesting that the driving parameters were not measured, or that too small of a sample set was collected to yield substantive correlations. This is not surprising given the complexity of the system and the large area over which the samples were collected.

Nonetheless, two notable observations were made. First, a comparison of the frequency of Firmicute 16S rRNA gene sequences obtained against the rate of methanogenesis in respective enrichments on coal suggests a correlation between the two parameters (Fig. 7). Second, rarefaction analysis of metabolite data with respect to the bacterial and archaeal 16S rRNA gene sequence data indicated that several metabolites were associated with increased bacterial and archaeal diversity (Fig. S4). A significantly greater diversity of archaeal OTUs was observed when 2-methylpentanoic acid, 5,6,7,8-tetrahydronaphthoic acid, butyric acid, methylnaphtoic acid, and *m*-tolylacetic acid were detected in produced water. Similarly, increased diversity of bacterial species was indicated in the presence of butyric acid, phenol, C7-unsaturated methylalkylsuccinate, and succinate.

Discussion

CBM methanogenesis pathway

Microbial CBM production was investigated across the over-pressured envelope of the SJB coal formation using enrichment studies as well as molecular characterization and metabolite profiling of produced waters. Integration of these data indicates the presence of a diverse community of methanogenic archaea, of which Methanomicrobiales and Methanosarcinales comprised the bulk of the detected core taxa. The order Methanosarcinales contains metabolically diverse methanogens capable of utilizing H_2/CO_2, acetate, and methyl compounds as substrates for methanogenesis (Whitman *et al.*, 2006). Specifically, sequence libraries from the SJB were enriched in two families within the Methanosarcinales (Fig. 5): the Methanosaeta and the Methanosarcina. Species within the family Methanosaeta are thought to be obligate acetate utilizers. Members of the family Methanosarcina also utilize acetate, but are also capable of more diverse metabolism that includes hydrogenotrophic methanogenesis as well as the conversion of methylated compounds to methane. The Methanomicrobiales were numerically less dominant in sequence libraries, but presented a greater richness of core taxa across the formation. All species within the order Methanomicrobiales are known to utilize hydrogen and carbon dioxide to generate methane, while none are capable of utilizing acetate (Garcia *et al.*, 2006). The order Methanomicrobiales also includes the

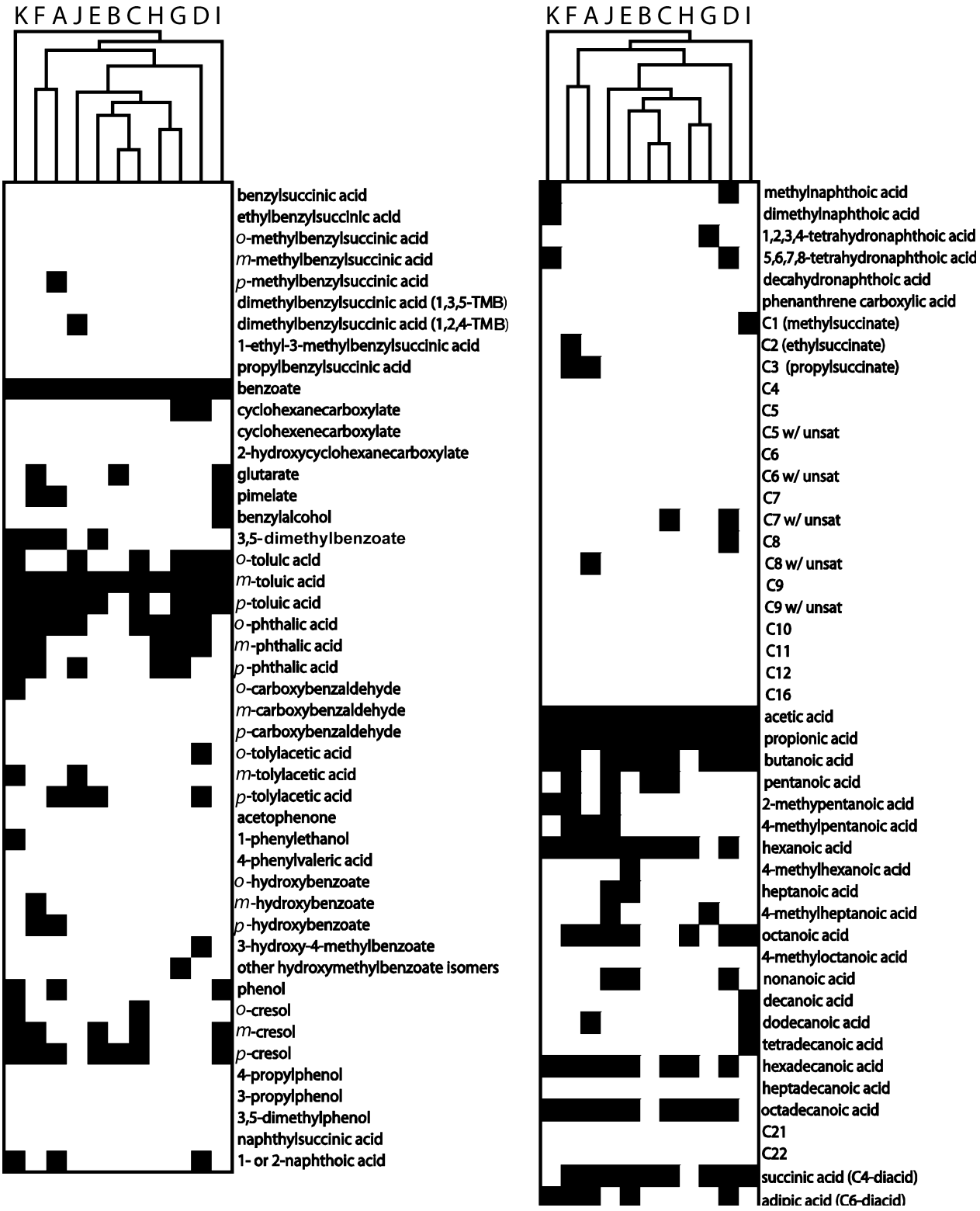

Fig. 6. Heat map demonstrating the relatedness of samples analyzed from the SJB based on the presence (black) or absence (white) of polar organic compounds that may be putative metabolites (to the right of heat map) implicated in anaerobic hydrocarbon/coal biodegradation. The samples interrogated by metabolite profiling are shown above the dendrogram. Hierarchical cluster analysis was performed in CLUSTER 3.0 using Euclidean distance measures. The heat map was created from the results of the cluster analysis and visualized in Java Tree View.

Table 2. GeoChip analysis: number of positive probes detected for individual biochemical pathways

	Sulfate reduction			Methanogenesis		Energy conservation		Anaerobic hydrocarbon activation and degradation							
Well	*aprA*	*dsrA*	*dsrB*	*mcrA*	*CODH*	*Hydrog.*	*Cyto.*	*assA*	*tutD*	*bbsG*	*ebdBC*	*bclA*	*bcr*	*badK*	*badH*
A	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
B	11	19	7	1	7	6	8	1	0	0	0	2	3	0	0
C	40	126	47	12	44	27	81	3	1	2	2	10	9	0	12
D	13	31	11	6	10	14	19	0	0	0	2	4	2	0	2
E	14	35	7	1	12	8	14	0	0	0	0	2	3	0	5
F	7	12	5	1	7	4	6	2	0	0	0	1	1	0	1
G	24	59	22	8	34	16	49	1	0	1	1	9	4	0	3
H	18	42	10	6	12	11	23	2	0	3	1	3	6	0	3
I	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
J	14	37	19	3	17	7	31	0	1	1	0	5	2	1	2
K	25	86	31	12	28	22	56	1	1	2	0	6	6	1	10

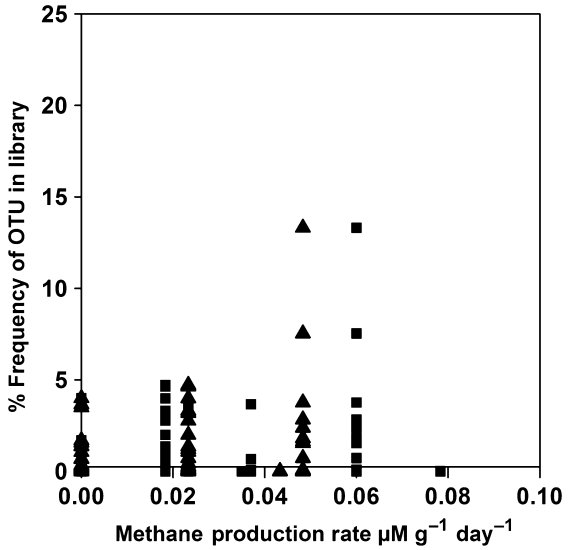

Fig. 7. Methanogenesis rates (μM g^{-1} day^{-1}) from enrichments using SJB coals and individual produced water samples vs. the frequency at which individual Firmicute OTUs were detected in sequence libraries generated from the inocula. Frequencies are shown as percentages of reads belonging to each OTU with respect to the total number of reads in each library. Symbols correspond to measurements from enrichments with two different coals: (▲) coal A R_o = 0.6% and (■) coal B/C R_o = 0.87%.

genus *Methanocorpusculum*. Microarray hybridization of community DNA (from all wells) to the *Methanocorpusculum mcrA* probe may indicate that this genus is widely distributed in the SJB (Fig. 3). However, hybridization to the *Methanocorpuculum* probe was also observed in well B, which had no PCR detectable archaea and produced no methane in enrichments. *Methanocorpusculum* was also not detected in the clone libraries from any of the sites, highlighting that caution should be exercised when assigning genera to functional gene sequences and that cross hybridization of probes cannot be excluded. Similarly, sequence frequencies within OTUs should not be given too much weight, given that their distribution can be heavily influenced by biases inherent in PCR, gene copy number per genome, and microbial growth rates.

Assuming that the physiology of cultivated methanogens is representative of methanogenic populations in the environment, these data suggest that populations within the Fruitland coal formation are methanogenic via both hydrogenotrophic and acetoclastic metabolism. The detected microbial diversity is consistent with previous studies reporting a range of methanogenic archaea associated with coals and their produced waters, including members of the *Methanosarcina*, *Methanolobus*, *Methanobacteria*, *Methanocorpusculum*, *Methanosaeta*, *Methanococci*, *Methanoculleus*, *and Methanoregula* genera (Strąpoć *et al.*, 2011). An analysis of the salient features of methanogenic populations across different formations is, unfortunately, difficult to achieve because of the diverse nature of prior sampling strategies and their limited scopes. Several notable exceptions include a study of natural gas producing wells in Cook Inlet, in which it was observed that *Methanolobus* and *Methanosarcina* contributed large proportions of the detected methanogenic archaeal sequences in libraries (Strąpoć *et al.*, 2010). Methanosarcinales accounted for the majority of the methanogens in the original coal samples from abandoned German mines (Beckmann *et al.*, 2011b) and also predominate in a consortium enriched from a CBM well from the Powder River Basin (Green *et al.*, 2008). Cultivated strains of these taxa can utilize methyl compounds, including methanol and methylamines, where *Methanolobus* is not known to utilize acetate or H_2/CO_2. These data, in addition to an analysis of δD CH_4, led the authors to conclude that biogenic methane in Cook Inlet

wells is primarily the result of acetoclastic and methylotrophic metabolism. This is in contrast to δD CH_4 data collected in the Illinois Basin, which suggested biogenic methane is produced primarily via the reduction of CO_2 (Strąpoć *et al.*, 2010). A limited phylogenetic survey of the Illinois Basin produced water and methanogenic enrichments indicated that the samples were almost exclusively dominated by the genus *Methanocorpusculum* (Strąpoć *et al.*, 2008), which is known to utilize H_2/CO_2, formate, and 2-propanol/CO_2 but not acetate. Overall, these data suggest that methanogenesis may proceed via different predominant mechanisms in different coal beds. Whether these differences are fortuitous or are rooted in the inherent geochemical characteristics of the requisite coals remains unclear.

Based on molecular profiling, the archaeal populations in the SJB likely have hybrid characteristics, containing taxa observed at both the Cook Inlet and the Illinois Basin sites. This interpretation is supported by the δD CH_4 measurements for the SJB compared with other sites (Strąpoć *et al.*, 2011). Carbon dioxide reduction in the Illinois basin is prevalent based on δD CH_4 values ranging from −200‰ to −220‰, while acetoclastic and methylotrophic methanogenesis predominates in Cook Inlet, yielding δD CH_4 values ranging between −280‰ and −300‰. Samples in the SJB exhibited δD CH_4 values between −210‰ and −235‰ (Strąpoć *et al.*, 2011), indicating that biogenic methane is mostly produced via CO_2 reduction, with acetoclastic and methylotrophic methanogenesis likely playing some role, but being of lesser importance. This observation is supported by enrichment studies (Table 1), which yielded little methane in the presence of acetate or propionate. High frequencies of Methanosaeta in sequence libraries, however, are not congruent with a lack of acetoclastic methanogenesis. Indeed, low molecular weight volatile fatty acids were readily detected in basin waters and presumably are both formed and consumed during the anaerobic biodegradation of coal constituents. Similarly, the presence of Methanosarcinales and the preferential utilization of methyl-containing substrates in the enrichments suggest the importance of methylotrophic methanogenesis, which is not supported by δD CH_4 measurements. It is unclear how to reconcile these observations. It may be that coal biodegradation processes are so tightly coupled that the exogenous addition of these substrates had little immediate impact. Alternatively, enrichment conditions may not have been favorable for the growth of the requisite acetoclastic methanogens. Other studies have certainly implicated acetate in CBM methane. Acetate was a major methane precursor using radiotracer techniques in coal slurry incubations and in field push–pull tests in the Powder River Basin (Ulrich & Bower, 2008). More recently, acetoclastic methanogenesis mediated by Methanosarcinales proved to be the predominate process in coal microcosms constructed from samples obtained from abandoned coal mines in Germany (Beckmann *et al.*, 2011a, b). However, discrepancies between geochemical and enrichment studies are not uncommon. For example, in enrichments of microorganisms associated with Powder River coal, it was demonstrated that acetate did not stimulate methane production (Harris *et al.*, 2008), even though δD CH_4 measurements tend to be in the same range as those observed for Cook Inlet (Strąpoć *et al.*, 2011). This highlights the need for more detailed molecular surveys of coals basins that exhibit a range of δD CH_4 values. Another caveat is the potential role of biofilms that may be associated with the surfaces of extraction wells. For example, *in situ* hydrogen release could result from metal being in contact with the production water. This in turn could enrich for hydrogenotrophic microorganisms (e.g. methanogens), and their differential detection in produced water may not be a reflection of their relative contribution in the formation *per se*. Sampling from freshly drilled wells might potentially minimize such a limitation.

Metabolite analysis

Irrespective of the specific methanogenic pathway, experiments described here clearly demonstrated that SJB coals represent a suitable carbon source for methanogenesis despite their relatively high thermal maturities and degree of aromatization. Nonetheless, no correlation between coal maturity and methanogenesis was observed. The absence of differences among samples of differing maturities might be explained by the shared history of the SJB coals. The long-term exposure of formation coals to resident microbial populations may have resulted in the equitable metabolism of the most labile coal constituents and their conversion to methane, while the more recalcitrant material remains, irrespective of the thermal maturity of the parent material. Unfortunately, this hypothesis is difficult to test because of the highly heterogeneous nature of coal, which does not lend itself to routine structural analysis.

Ultimately, though, coal must be broken down into simpler chemical residues, some of which are likely amenable to GC–MS analysis. An assay for putative metabolites suggested the predominance of aromatic compound utilization (benzoate, toluic acids, phthalic acids, and cresols), and functional gene analysis corroborated this contention by providing evidence for the presence of genes consistent with anaerobic biodegradation of aromatic molecules. Hence, it appears that the parent substrate (i.e. a subset of the coal matrix) is broken down into aromatic intermediates that are further metabolized

via fermentative metabolism, eventually yielding methanogenesis precursors. This hypothesis is supported by rarefaction analysis of archaeal 16S sequence data, which suggests that a significant association exists between archaeal diversity and individual aromatic or short-chain fatty acid metabolites (Fig. S4). *m*-Tolylacetic acid, 5,6,7,8-tetrahydronaphthoic acid, and methylnaphthoic acid are likely aromatic breakdown products of bacterial coal biodegradation. In the absence of alternate electron acceptors, these substrates are likely further metabolized into short-chain fatty acids such as butyric or 2-methylpentanoic acid that serve as the basis for a variety of syntrophic interactions that stimulate the requisite methanogens. Enrichment experiments also indicate that fermentative production of short-chain alcohols (methanol, ethanol, propanol) may play a role in the community metabolism of coal (Table 2), but these substrates escape our routine extraction and GC–MS analysis procedures.

In addition to aromatic ring degradation, it has been suggested that *n*-alkanes may play a role in coal-bed methanogenesis (Scott *et al.*, 1994). Over large parts of the north-central area of the SJB, coals of higher rank are significantly depleted in *n*-alkanes (Scott *et al.*, 1994). Thermal maturation and gas solution stripping have been invoked as explanations for this pattern of distribution but are not supported by the presumed geological history of the formation (Scott *et al.*, 1994). Data presented here indicate that *n*-alkane degrading bacteria are present in the SJB production waters and that metabolites consistent with anaerobic alkane metabolism are produced. The degree to which *n*-alkane degradation contributes to methane production in coals is unknown, but it appears to be relatively widely distributed, at least in SJB coals. Nine of eleven tested samples either contained alkylsuccinate metabolites or exhibited hybridization to at least one of the *assA* gene probes represented on the GeoChip.

Bacterial communities associated with CBM

Undoubtedly, coal methanogenesis requires a complex community of bacterial and archaeal species that cooperatively degrade the complex coal matrix via a sequence of fermentation and syntrophic interactions. A basin-wide analysis of bacterial 16S rRNA genes in the SJB generally supports this notion (Fig. 5). Dominant among the bacterial species are the Proteobacteria and Firmicutes, which contain many well-known fermentative organisms (particularly the Clostridiales). Among the detected proteobacterial lineages, the deltaproteobacteria comprised the largest fraction and included taxa belonging to the Syntrophobacterales as well as several sulfate-reducing bacterial groups known to participate in syntrophic relationships with methanogens in the absence of sulfate. This is certainly consistent with the low levels of sulfate within the tested regions of the SJB. Comparable data are not available for other coal formations, and an evaluation of their salient cross-basin features can therefore not be made. In a limited study of 16S rRNA genes from Illinois Basin production waters, alphaproteobacteria, Firmicutes, Bacteroidetes, and Spirochetes were detected, but the sample size in this study was small (Strąpoć *et al.*, 2008). A more detailed study of bacterial populations in a coal seam aquifer located in northern Japan demonstrated that populations were dominated by *Acetobacterium* and *Syntrophus* (Shimizu *et al.*, 2007), implicating syntrophic utilization of some coal degradation products. Syntrophic consortia were also obtained in enrichments of well-bore water from the Powder River basin (Green *et al.*, 2008). These data are consistent with observations of bacterial communities in the SJB, where 16S signatures of known syntrophic bacteria were widespread. Overall, these data suggest that syntrophic metabolism is an important and general feature of biogenic CBM formation.

Syntrophic metabolism represents only the final stages in CBM production. A distinct and different component of the bacterial community is likely to be involved in the depolymerization and activation of the coal matrix, a process that is thought to represent the rate-limiting step of biogenic CBM formation. The dominant signals among the SJB core taxa were Firmicute and Actinomycete lineages. These taxonomic groups are known for their versatile metabolic activity. Actinomycetes are particularly important in soil environments where they participate in the biodegradation of lignocellulosic materials. Their role in anaerobic coal formations has not been recognized thus far, but data presented here point to their potential significance. More importantly, Firmicute lineages belonging to the Clostridiales and Thermoanaerobacterales represent the largest fraction of core taxa across the SJB. The Clostridiales are a diverse group of metabolically versatile, spore-forming, anaerobic bacteria that can catalyze a wide range of metabolic bio-conversions. Clostridia are known to depolymerize starch, chitin, xylan, and cellulose and are known to occur in diverse anaerobic habitats such as the rumen, soils, and sediments (Wiegel *et al.*, 2006). Similarly, the Thermoanaerobacterales include thermophilic, anaerobic, fermentative bacteria capable of utilizing a variety of carbon substrates including simple and complex carbohydrates. Their importance has been recognized for geothermally heated environments and high-temperature petroleum reservoirs (Wiegel, 2009), suggesting their importance in hydrocarbon-bearing formations. Here, it was observed that the frequency of Firmicute 16S rRNA gene sequences in produced water may correlate with methanogenesis rates in enrichments on coal (Fig. 7). While these data do not conclusively implicate Firmicutes

in coal activation, they do represent an interesting starting point by providing a hypothesis (i.e. that Firmicutes may play an important role). Firmicutes have previously been recognized to be important components of coal-bed communities, when samples from an Australian CBM reservoir were found to contain an abundance of lineages within the Clostridiales (Li *et al.*, 2008).

Conclusion

Overall, it was confirmed that coals in the SJB can be converted to methane and that both aromatic and aliphatic components serve as substrates that ultimately lead to methane formation. SJB production waters were observed to contain diverse bacterial and archaeal communities that can interact in important ways to ultimately catalyze methane formation. While most emphasis in the literature has been placed on the nature of the methanogens involved in CBM formation, data presented here suggest that this emphasis is somewhat misplaced. Our findings suggest that it is unlikely that the availability of both acetoclastic and hydrogenotrophic methanogens will inherently limit coal methanogenesis. Rather, more emphasis should be placed on metabolic steps that serve to initiate anaerobic coal biodegradation processes. In this regard, the identity of bacterial community members responsible for the anaerobic activation of the coal matrix carbon remains elusive, but correlation analysis makes it tempting to hypothesize that Firmicute lineages may play an important role. In this context, a renewed focus on the factors that inherently limit coal methanogenesis, particularly in the context of the bacterial community, will allow for more strategic attempts to stimulate *in situ* production.

Acknowledgements

This work was funded by in part by ConocoPhillips Corporation, Houston, TX, and in part through a grant from the National Science Foundation (Grant Number: MCB-0921265). We would like to thank Ken Walker from Champion Technologies, Inc. (Farmington NM) for support in the field. We would also like to acknowledge and thank Annie Frost at ConocoPhillips Corporation for field assistance. The helpful advice and assistance of Brian J. Cardott in obtaining coals from the Oklahoma Geological Survey collection are gratefully acknowledged.

References

Abu Laban N, Selesi D, Jobelius C & Meckenstock RU (2009) Anaerobic benzene degradation by Gram-positive sulfate-reducing bacteria. *FEMS Microbiol Ecol* **68**: 300–311.

Abu Laban N, Selesi D, Rattei T, Tischler P & Meckenstock RU (2010) Identification of enzymes involved in anaerobic benzene degradation by a strictly anaerobic iron-reducing enrichment culture. *Environ Microbiol* **12**: 2783–2796.

Beckmann S, Kruger M, Engelen B, Grobushina A & Cypionka H (2011a) Role of bacteria, archaea and fungi involved in methane release in abandoned coal mines. *Geomicrobiol J* **28**: 347–358.

Beckmann S, Lueders T, Kruger M, von Netzer F, Engelen B & Cypionka H (2011b) Acetogens and acetoclastic methanosarcinales govern methane formation in abandoned coal mines. *Appl Environ Microbiol* **77**: 3749–3756.

Beller HR, Kane SR, Legler TC, McKelvie JR, Lollar BS, Pearson F, Balser L & Mackay DM (2008) Comparative assessments of benzene, toluene, and xylene natural attenuation by quantitative polymerase chain reaction analysis of a catabolic gene, signature metabolites, and compound-specific isotope analysis. *Environ Sci Technol* **42**: 6065–6072.

Bergmann FD, Selesi D & Meckenstock RU (2011) Identification of new enzymes potentially involved in anaerobic naphthalene degradation by the sulfate-reducing enrichment culture N47. *Arch Microbiol* **193**: 241–250.

Boll M & Heider J (2010) Anaerobic degradation of hydrocarbons: mechanisms of C–H-bond activation in the absence of oxygen. *Handbook of Hydrocarbon and Lipid Microbiology* (Timmis KN, McGenity T, Van der Meer JR & De Lorenzo V, eds), pp. 1011–1024. Springer-Verlag, Berlin, Heidelberg.

Callaghan AV, Wawrik B, Ni Chadhain SM, Young LY & Zylstra GJ (2008) Anaerobic alkane-degrading strain AK-01 contains two alkylsuccinate synthase genes. *Biochem Biophys Res Commun* **366**: 142–148.

Callaghan AV, Davidova IA, Savage-Ashlock K *et al.* (2010) Diversity of benzyl- and alkylsuccinate synthase genes in hydrocarbon-impacted environments and enrichment cultures. *Environ Sci Technol* **44**: 7287–7294.

Carmona M, Zamarro MT, Blazquez B *et al.* (2009) Anaerobic catabolism of aromatic compounds: a genetic and genomic view. *Microbiol Mol Biol Rev* **73**: 71–133.

Duncan KE, Gieg LM, Parisi VA, Tanner RS, Tringe SG, Bristow J & Suflita JM (2009) Biocorrosive thermophilic microbial communities in Alaskan North Slope oil facilities. *Environ Sci Technol* **43**: 7977–7984.

Elshahed MS, Gieg LM, McInerney MJ & Suflita JM (2001) Signature metabolites attesting to the *in situ* attenuation of alkylbenzenes in anaerobic environments. *Environ Sci Technol* **35**: 682–689.

EPA US (2004) Evaluation of impacts to underground sources of drinking water by hydraulic fracturing of coalbed methane reservoirs study. Attachment 1 The San Juan Basin: EPA 816-R-804-003.

Erkel C, Kube M, Reinhardt R & Liesack W (2006) Genome of Rice Cluster I archaea – the key methane producers in the rice rhizosphere. *Science* **313**: 370–372.

Faiz M & Hendry P (2006) Significance of microbial activity in Australian coal bed methane reservoirs – a review. *Bull Can Petrol Geol* **54**: 261–272.

Fry JC, Horsfield B, Sykes R *et al.* (2009) Prokaryotic populations and activities in an interbedded coal deposit, including a previously deeply buried section (1.6–2.3 km) above 150 Ma basement rock. *Geomicrobiol J* **26**: 163–178.

Garcia J-L, Ollivier B & Whitman WB (2006) The order Methanomicrobiales. *The Prokaryotes* (Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E & Dworkin M, eds), pp. 208–230. Springer Verlag, New York.

Gieg LM & Suflita JM (2002) Detection of anaerobic metabolites of saturated and aromatic hydrocarbons in petroleum-contaminated aquifers. *Environ Sci Technol* **36**: 3755–3762.

Gieg LM, Parisi VA & Suflita JM (2009) Metabolic indicators of anaerobic hydrocarbon decay in diverse petroleum-laden environments. *Geochim Cosmochim Acta* **73**: A434.

Green MS, Flanegan KC & Gilcrease PC (2008) Characterization of a methanogenic consortium enriched from a coalbed methane well in the Powder River Basin, USA. *Int J Coal Geol* **76**: 34–45.

Grundmann O, Behrends A, Rabus R, Amann J, Halder T, Heider J & Widdel F (2008) Genes encoding the candidate enzyme for anaerobic activation of n-alkanes in the denitrifying bacterium, strain HxN1. *Environ Microbiol* **10**: 376–385.

Hamady M, Walker JJ, Harris JK, Gold NJ & Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nat Methods* **5**: 235–237.

Harris SH, Smith RL & Barker CE (2008) Microbial and chemical factors influencing methane production in laboratory incubations of low-rank subsurface coals. *Int J Coal Geol* **v. 76**: 46–51.

He Z & Zhou J (2008) Empirical evaluation of a new method for calculating signal-to-noise ratio for microarray data analysis. *Appl Environ Microbiol* **74**: 2957–2966.

He ZL, Deng Y, Van Nostrand JD *et al.* (2010) GeoChip 3.0 as a high-throughput tool for analyzing microbial community composition, structure and functional activity. *ISME J* **4**: 1167–1179.

de Hoon MJ, Imoto S, Nolan J & Miyano S (2004) Open source clustering software. *Bioinformatics* **20**: 1453–1454.

Johnson HA, Pelletier DA & Spormann AM (2001) Isolation and characterization of anaerobic ethylbenzene dehydrogenase, a novel Mo-Fe-S enzyme. *J Bacteriol* **183**: 4536–4542.

Klein DA, Flores RM, Venot C *et al.* (2008) Molecular sequences derived from Paleocene Fort Union Formation coals vs. associated produced waters: implications for CBM regeneration. *Int J Coal Geol* **76**: 3–13.

Kniemeyer O & Heider J (2001) Ethylbenzene dehydrogenase, a novel hydrocarbon-oxidizing molybdenum/iron-sulfur/heme enzyme. *J Biol Chem* **276**: 21381–21386.

Kniemeyer O, Musat F, Sievert SM *et al.* (2007) Anaerobic oxidation of short-chain hydrocarbons by marine sulphate-reducing bacteria. *Nature* **449**: 898–901.

Krüger M, Beckmann S, Engelen B, Thielemann T, Cramer B, Schippers A & Cypionka H (2008) Microbial methane formation from hard coal and timber in an abandoned coal mine. *Geomicrobiol J* **25**: 315–321.

Li DM, Hendry P & Faiz M (2008) A survey of the microbial populations in some Australian coalbed methane reservoirs. *Int J Coal Geol* **76**: 14–24.

Ludwig W, Strunk O, Westram R *et al.* (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* **32**: 1363–1371.

McInerney MJ, Bryant MP & Pfennig N (1979) Anaerobic bacterium that degrades fatty-acids in syntrophic association with methanogens. *Arch Microbiol* **122**: 129–135.

Musat F, Galushko A, Jacob J *et al.* (2009) Anaerobic degradation of naphthalene and 2-methylnaphthalene by strains of marine sulfate-reducing bacteria. *Environ Microbiol* **11**: 209–219.

Parisi VA, Brubaker GR, Zenker MJ *et al.* (2009) Field metabolomics and laboratory assessments of anaerobic intrinsic remediation of hydrocarbons at a petroleum-contaminated site. *Microb Biotechnol* **2**: 202–212.

Penner TJ, Foght JM & Budwill K (2010) Microbial diversity of western Canadian subsurface coal beds and methanogenic coal enrichment cultures. *Int J Coal Geol* **82**: 81–93.

Pérez-Pantoja D, González B & Pieper DH (2010) Aerobic degradation of aromatic hydrocarbons. *Handbook of Hydrocarbon and Lipid Microbiology* (Timmis KN, McGenity T, Meer JRVD & Lorenzo VD, eds), pp. 799–837. Springer Verlag, Berlin.

Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J & Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res* **35**: 7188–7196.

Rainey FA, WardRainey N, Kroppenstedt RM & Stackebrandt E (1996) The genus *Nocardiopsis* represents a phylogenetically coherent taxon and a distinct actinomycete lineage: proposal of Nocardiopsaceae fam nov. *Int J Syst Bacteriol* **46**: 1088–1092.

Saldanha AJ (2004) Java Treeview – extensible visualization of microarray data. *Bioinformatics* **20**: 3246–3248.

Schloss PD, Westcott SL, Ryabin T *et al.* (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* **75**: 7537–7541.

Scott AR, Kaiser WR & Ayers WB (1994) Thermogenic and secondary biogenic gases, San-Juan Basin, Colorado and New-Mexico – implications for coalbed gas producibility. *AAPG Bull* **78**: 1186–1209.

Shimizu S, Akiyama M, Naganuma T, Fujioka M, Nako M & Ishijima Y (2007) Molecular characterization of microbial communities in deep coal seam groundwater of northern Japan. *Geobiology* **5**: 423–433.

Strąpoć D, Picardal FW, Turich C *et al.* (2008) Methane-producing microbial community in a coal bed of the Illinois basin. *Appl Environ Microbiol* **74**: 2424–2432.

Strąpoć D, Ashby M, Wood L, Levinson R & Huizinga B (2010) Significant contribution of methyl/methanol-utilizing

methanogenic pathway in a subsurface biogas environment. *Applied Microbiology and Molecular Biology in Oil Field Systems* (Whitby C & Skovhus TL, eds), pp. 250. Springer, Berlin.

Strąpoć D, Mastalerz M, Dawson K *et al.* (2011) Biogeochemistry of microbial coal-bed methane. *Annu Rev Earth Planet Sci* **38**: 617–656.

Taylor BF, Campbell WL & Chinoy I (1970) Anaerobic degradation of the benzene nucleus by a facultatively anaerobic microorganism. *J Bacteriol* **102**: 430–437.

Thielemann T, Cramer B & Schippers A (2004) Coalbed methane in the Ruhr Basin, Germany: a renewable energy source? *Org Geochem* **34**: 1537–1549.

Tierney M & Young LY (2010) Anaerobic degradation of aromatic hydrocarbons. *Handbook of Hydrocarbon and Lipid Microbiology* (Timmis KN, McGenity T, Van der Meer JR & De Lorenzo V, eds), pp. 925–934. Springer-Verlag, Berlin, Heidelberg.

Ulrich G & Bower S (2008) Active methanogenesis and acetate utilization in Powder River Basin Coals. *Int J Coal Geol* **76**: 25–33.

Van Nostrand JD, Wu WM, Wu L *et al.* (2009) GeoChip-based analysis of functional microbial communities during the reoxidation of a bioreduced uranium-contaminated aquifer. *Environ Microbiol* **11**: 2611–2626.

Wang Q, Garrity GM, Tiedje JM & Cole JR (2007) Naive Bayesïan classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* **73**: 5261–5267.

Whitman WB, Bowen TL & Boone DR (2006) The methanogenic bacteria. *The Prokaryotes*, Vol. **3** (Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E & Dworkin M, eds), pp. 165–207. Springer Verlag, New York.

Widdel F & Bak F (1992) Gram-negative mesophilic sulfate-reducing bacteria. *The Prokaryotes* (Balows A, Trüper HG, Dworkin M, Harder W & Schleifer K-H, eds), pp. 3352–3378. Springer, New York, NY.

Widdel F & Grundmann O (2010) Biochemistry of the anaerobic degradation of non-methane alkanes. *Handbook of Hydrocarbon and Lipid Microbiology* (Timmis KN, McGenity T, Van der Meer JR & De Lorenzo V, eds.), pp. 909–924. Springer-Verlag, Berlin, Heidelberg.

Widdel F & Musat F (2010) Diversity and common principles in enzymatic activation of hydrocarbons. *Handbook of Hydrocarbon and Lipid Microbiology* (Timmis KN, McGenity T, Van der Meer JR & De Lorenzo V, eds), pp. 981–1009. Springer-Verlag, Berlin, Heidelberg.

Wiegel J (2009) Order III. Thermoanaerobacterales ord. nov. *Bergey's Manual of Systematic Bacteriology, Vol. 3: The Firmicutes* (Whitman WB, ed), p. 1224. Springer Verlag, New York, NY.

Wiegel J, Tanner R & Rainey FA (2006) Volume 4: bacteria: firmicutes, cyanobacteria. *The Prokaryotes* (Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E & Dworkin M, eds), pp. 654–678. Springer Verlag, New York, NY.

Wu L, Liu X, Schadt CW & Zhou J (2006) Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. *Appl Environ Microbiol* **72**: 4931–4941.

Yagi JM, Suflita JM, Gieg LM, DeRito CM, Jeon CO & Madsen EL (2010) Subsurface cycling of nitrogen and anaerobic aromatic hydrocarbon biodegradation revealed by nucleic Acid and metabolic biomarkers. *Appl Environ Microbiol* **76**: 3124–3134.

Zhou Z, Ballentine CJ, Kipfer R, Schoell M & Thibodeaux S (2005) Noble gas tracing of groundwater/coalbed methane interaction in the San Juan Basin, USA. *Geochim Cosmochim Acta* **69**: 5413–5428.

Supporting Information

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

Fig. S1. Rarefaction analysis of archaeal 16S gene sequences.

Fig. S2. Heatmap of the distribution of archaeal core taxa.

Fig. S3. Class-level phylogenetic breakdown of bacterial 16S rRNA gene sequences detected in each of the San Juan Basin produced water samples.

Fig. S4. Rarefaction analysis of 16S data with respect to detected metabolites.

Fig. S5. Ordination plot using non-metric multidimensional scaling.

Table S1. pH and solute concentrations of produced waters.

Please note: Wiley-Blackwell is 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.