Supplementary Information

Comparative analysis of metagenomes from three methanogenic hydrocarbon-degrading enrichment cultures with 41 environmental samples

Boon-Fei Tan[†1a], S. Jane Fowler[†2b], Nidal Abu Laban[1,2], Xiaoli Dong[3c], Christoph W. Sensen[3d], Julia Foght[1], Lisa M. Gieg[2*]

***Corresponding author** Tel.: 403-210-7207; FAX: 403-289-9311; e-mail: lmgieg@ucalgary.ca

[1]Department of Biological Sciences, University of Alberta, 116 St. and 85 Ave, Edmonton, Alberta, Canada T6G 2E9

[2]Department of Biological Sciences, University of Calgary, 2500 University Dr. NW, Calgary, Alberta, Canada T2N 1N4

[3]Visual Genomics Centre, Faculty of Medicine, 3330 Hospital Drive NW, Calgary, Alberta, Canada, T2N 4N1

[a]Present address: Singapore-MIT Alliance for Research and Technology, 1 Create Way, #09-03 Create Tower, Singapore 138602

[b]Present address: Department of Environmental Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

[c]Present address: Department of Geoscience, University of Calgary, 2500 University Dr. NW, Calgary, Alberta, Canada T2N 1N4

[d]Present address: Graz University of Technology, Institute of Molecular Biotechnology, Petersgasse 14, A-8010 Graz, Austria

[†]These authors contributed equally to this work

Contains:

Supplementary Materials and Methods:

NAPDC cultivation (anaerobic naphtha-degrading enrichment culture)

Growth conditions:

Methanogenic naphtha-degrading enrichment cultures were established using mature fine tailings (MFT) collected from Mildred Lake Settling Basin (MLSB; 57.057°N 111.600°W), an oil sands tailings pond in northeastern Alberta, Canada. Duplicate parent cultures were established by mixing 50 mL of MFT and 50 mL anoxic medium in 158 mL serum bottles with a headspace of 80% N_2, balance CO_2 (Siddique *et al*., 2007). The medium used for the enrichment was basal mineral fresh water medium (Widdel and Bak, 1992). The parent cultures were amended with 0.2% v/v naphtha (CAS no. 64742-49-0) as the hydrocarbon substrate. After 365 days of the incubation, the parent cultures produced 0.9 mmol CH_4.

After incubation for 372 d with intermittent monitoring for methane production, the parent cultures were pooled together and transferred as a 10 vol% inoculum into twelve serum bottles containing the same sterile anaerobic basal fresh water medium. These cultures again received the 0.2% v/v naphtha as the hydrocarbon substrate. Duplicate bottles without amendments were used as endogenous controls (substrate-unamended) to account for methane produced from residual hydrocarbons. Other duplicate bottles contained medium and 0.2% v/v naphtha only to account for any abiotic loss of hydrocarbons. Cultures were monitored for CH_4 production and hydrocarbon loss (see below). After 103 days of the incubation, the 10 bottles of the active naphtha-degrading cultures were sacrificed for metagenome analysis.

Analytical procedures:

Methane measurements were performed by mixing the culture bottles by hand and removing 100 µl of headspace gas using a 0.5-mL 28-gauge insulin syringes (Becton Dickinson, Franklin Lakes NJ). Samples were injected in a Hewlett Packard 5700A GC equipped with a flame ionization detector (FID) and analyzed as previously described (Holowenko *et al*., 2000).

For analysis of naphtha, 100 µL of headspace was analyzed for the presence of volatile hydrocarbons by direct injection into an Agilent 6890N GC with a model 5973 inert mass selective detector, fitted

with an Agilent capillary column (HP-5MS, 30 m in length, 0.25 mm internal diameter, 0.25 μm film thickness). The injector temperature was 250°C, with helium as carrier gas at a flow rate of 1 mL min-1. The oven temperature program was 40°C for 5 min, 10 °C min-1 to 85°C, and finally 5 min held at 85°C. Based on this analysis method, 45 peaks of hydrocarbons were identified (Figure S1). *m*- and *p*-xylene were not completely resolved under the GC conditions used, therefore the sum of their peaks in the mixture is reported.

Supplementary Results and Discussion

Methanogenesis pathways

The presence of genes encoding pathways for hydrogenotrophic, acetotrophic, and methylotrophic methanogenesis (Figure S4A) was examined using Pfam models. We found hits for all relevant models for hydrogenotrophic and acetotrophic methanogenesis in the three metagenomes (Figure S4B). However, Pfam hits to MtaB (pf12176), involved in methylotrophic methanogenesis, were not found in any of the cultures (Figure S4B), suggesting that methylotrophic methanogenesis either does not occur or is rare in these cultures and that this gene escaped sequencing. To ensure that the taxonomic affiliations of Pfam hits were appropriate for methanogenesis genes, the contigs containing sequences of interest were compared to the NCBI nr database using BLAST. Pfam recovered more sequences related to diverse methanogenesis pathways in SCADC and TOLDC including hydrogenotrophic pathways in *Methanoculleus* and *Methanoregula*, and acetoclastic methanogenesis in *Methanosaeta* (Figure S4B). Acetoclastic pathways in *Methanosaeta* were less prominent in NAPDC although the organism is more abundant based on 16S pyrotag sequencing. This is likely due to the relatively low sequencing coverage in NAPDC, or primer or sequencing bias in 16S rRNA gene pyrotag sequencing. Evidently, mapping of metagenomic reads to selected reference genomes of methanogens showed that NAPDC, SCADC, and TOLDC all harbour a strain of *Methanosaeta* that is closely related to *Methanosaeta concilli* GP6, though NAPDC had a considerably lower coverage. The *Methanoculleus* detected in these cultures is not similar to sequenced members of this genus, with ≤ 2.2% of the genome being mapped to the *Methanoculleus marisnigri* JR1 reference genome (Table S2).

Other metabolic pathways

We examined the three enrichment culture metagenomes for the presence of other pathways involved in anaerobic metabolism using the KEGG module mapper. Briefly, KEGG orthologues annotated in MG-RAST (Meyer *et al.*, 2008) as described in the main text were used as input to KEGG mapper to visualize metabolic pathways of interest (Kanehisa *et al.*, 2012).

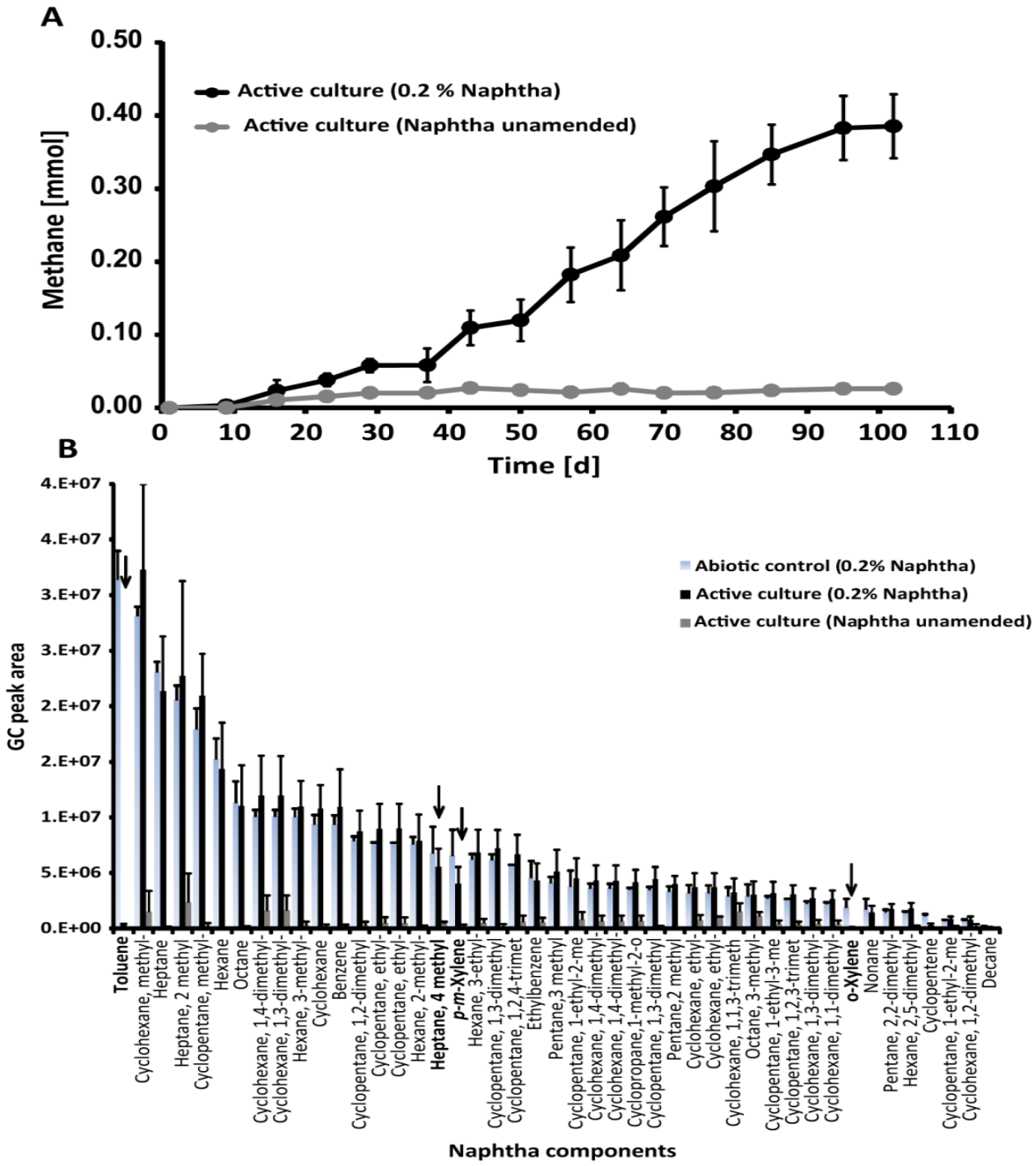
Pathways of interest included fumarate regeneration, β-oxidation, carbon fixation, the glyoxylate cycle, and alternate electron-accepting pathways (Figure S5, Table S4). Fumarate regeneration from succinyl-CoA (via key enzymes encoded by *bbsEF* and *sdh*) plays an important role in aromatic hydrocarbon degradation by addition to fumarate (Leutwein and Heider, 1999). In the cultures studied here, both *bbsEF* and *sdh* were identified, supporting the proposal that fumarate addition and recycling occurs in the cultures. During *n*-alkane degradation under sulfate-reducing conditions, fumarate is presumably regenerated from propionyl-CoA via the methylmalonyl-CoA pathway (Davidova *et al.*, 2005). In NAPDC and TOLDC, methylmalonyl-CoA racemase genes were not detected (possibly due to incomplete sequencing), but genes encoding the complete methylmalonyl-CoA pathway were found located close to a putative *ass* operon in the SCADC metagenome (Tan *et al*., 2013). Downstream alkane metabolism proceeds via β-oxidation and complete pathways for β-oxidation were found in all three methanogenic cultures (Figure S5).

Further oxidation of acetate from β−oxidation could occur via the oxidative TCA cycle. The requisite genes for a complete TCA cycle were found in all three cultures, which was surprising since this pathway is often incomplete in strict anaerobes. Some denitrifying toluene degraders are able to assimilate acetyl-CoA from toluene via the glyoxylate cycle, allowing carbon assimilation to occur solely from the hydrocarbon substrate (Heider and Fuchs, 1997). Although the glyoxylate cycle is often not present in strict anaerobes (Heider and Fuchs, 1997) the requisite genes were found in TOLDC (Figure S5). In strict anaerobes, carbon for biosynthesis cannot be entirely derived from the hydrocarbon substrate, and significant levels of CO_2 fixation concomitant with hydrocarbon metabolism have been observed in anaerobic hydrocarbon-degrading cultures (Taubert *et al.*, 2012; Winderl *et al.*, 2010). Carbon fixation pathways such as the reductive TCA cycle, Wood-Ljungdahl and/or 3-hydroxypropionate pathway may be active. Interestingly, all three cultures contained pyruvate synthase (reductive TCA cycle) and so conceivably could assimilate carbon by carboxylation of acetyl-CoA (Figure S5). This could then lead into the reductive TCA cycle for further carbon assimilation, which was fully annotated in all three of the cultures (Figure S5). Several genes for the Wood-

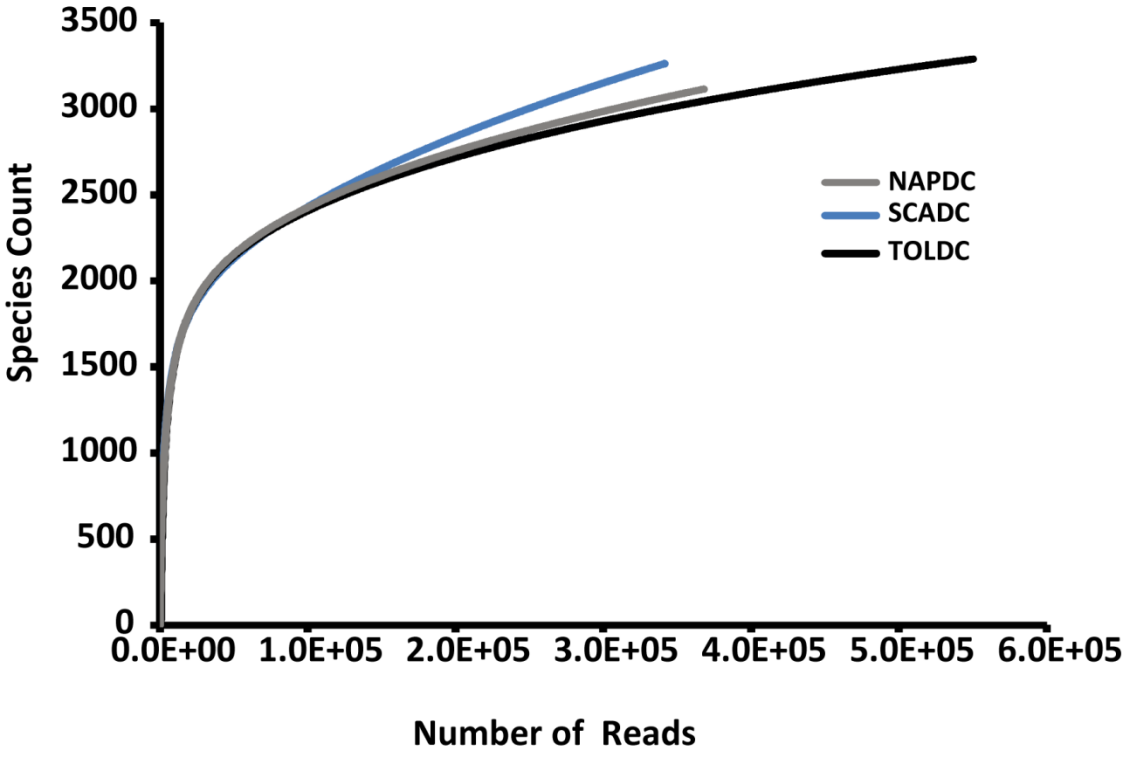
Ljungdahl and 3-hydroxypropionate pathways were not detected in TOLDC and NAPDC, but were present in SCADC (Figure S5), showing that this culture possesses multiple carbon fixation pathways. Therefore, the reductive TCA cycle may be a primary mode of carbon fixation in TOLDC and NAPDC.

Because all three methanogenic cultures were initially enriched from environments containing alternate electron acceptors, we also searched their metagenomes for canonical genes associated with dissimilatory nitrate and sulfate reduction pathways (Figure S5). Only incomplete pathways for nitrate reduction were annotated in the metagenomes. Due to the low abundance of known nitrate-reducing bacteria in the methanogenic cultures (Table 3, main text), the absence of complete nitrate reduction operons in the metagenomes is not unexpected. However, incomplete dissimilatory nitrate reduction may be possible. The pathway for sulfate reduction was identified in NAPDC and SCADC, but the key gene *dsrA* was not identified in TOLDC, suggesting either that the toluene-degrading culture is incapable of dissimilatory sulfate reduction, or that *dsrA* was present at sufficiently low abundance to escape metagenomic sequencing. Although the TOLDC bacterial community was dominated by fermentative and/or syntrophic Firmicutes, there was a relatively high abundance of putative sulfate-reducers present in SCADC and particularly in NAPDC (Table 3, main text), which are 'younger' methanogenic enrichments than TOLDC and may not yet have lost sulfate-reducing members.

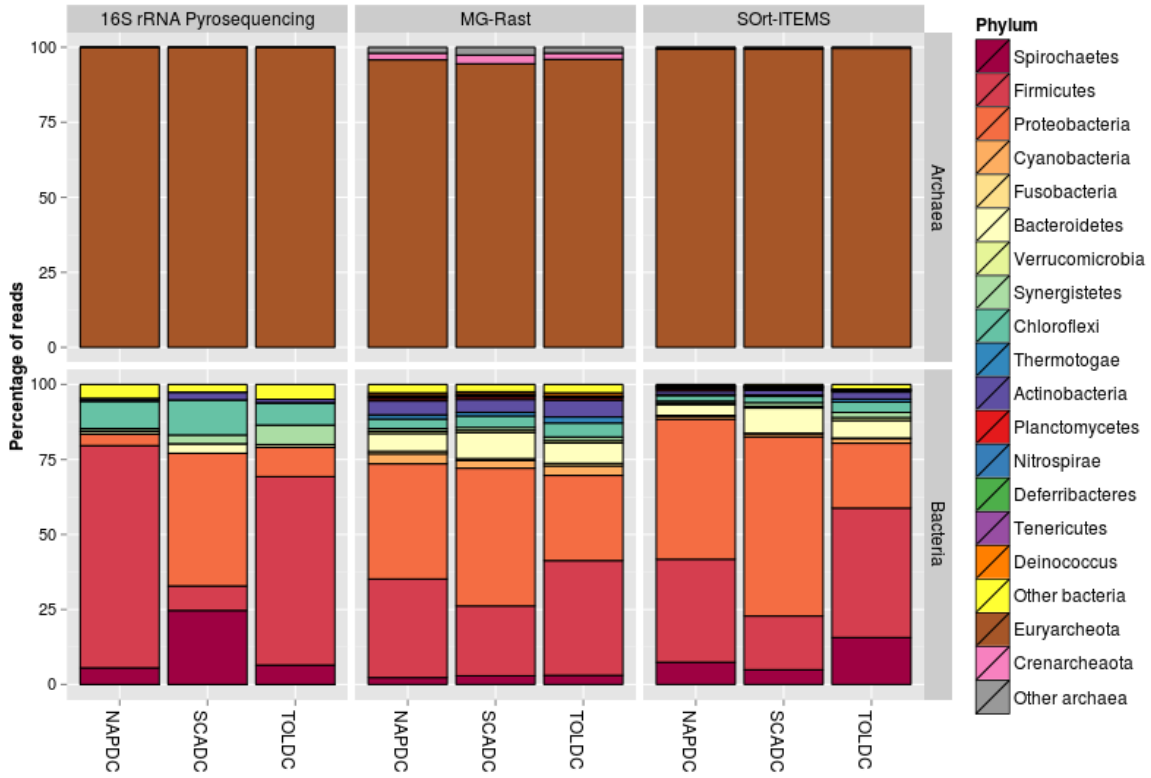
Supplementary Figures



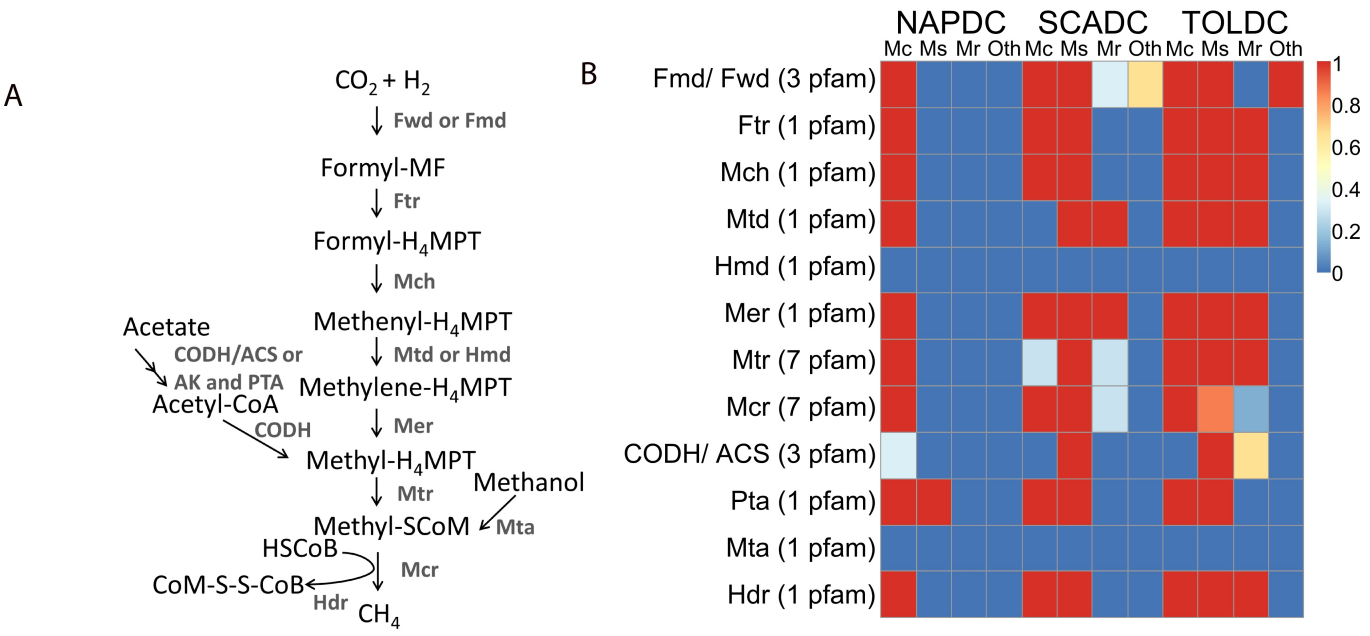
Supplementary Figure S1. (A) Methane generation from NAPDC incubated with naphtha during a 103-d incubation (a total of 0.4 mmol CH_4 was produced), and **(B)** degradation of aromatic and aliphatic hydrocarbon components (indicated in boldface with arrows) at the end of the 103-d incubation in the NAPDC inoculated cultures and controls. Changes in GC peak areas were monitored in cultures (black bars) compared to abiotic controls (blue bars). The analytical method used for hydrocarbon analysis was able to resolve 45 hydrocarbon compounds within naphtha. The complete degradation of toluene and *o*-xylene and partial degradation of *m*- and *p*-xylene and 4-methylheptane after the 103 d incubation of NAPDC was observed relative to controls.



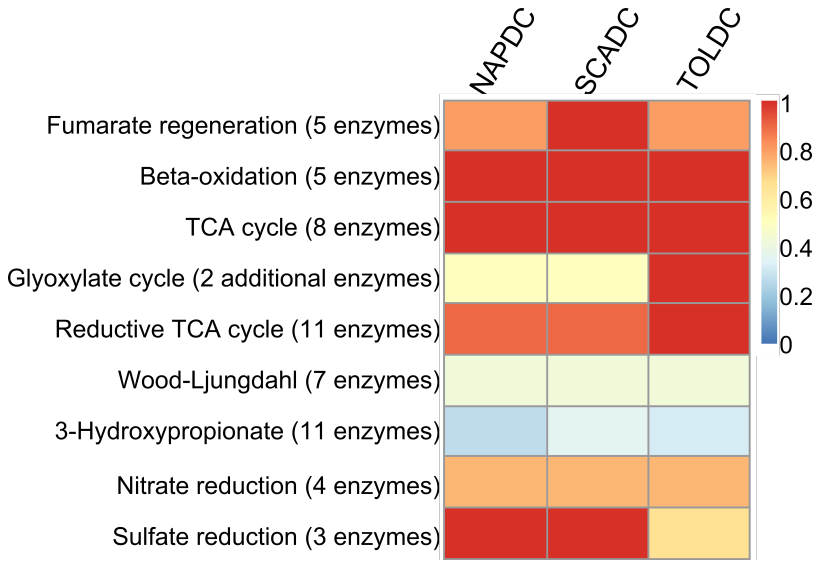
Supplementary Figure S2. Rarefaction curve of annotated species richness for NAPDC, SCADC, and TOLDC reflecting the total number of distinct species annotations as a function of the number of metagenome reads.



Supplementary Figure S3. Composition of microbial communities in NAPDC, SCADC, and TOLDC determined from 16S rRNA gene amplicon (pyro)sequences or unassembled 454 metagenomic reads accessed using MG-Rast and SOrt-ITEMS. Relative abundance is expressed as the proportion of total archaeal reads (top panel) or bacterial reads (bottom panel).



Supplementary Figure S4. Enzymes involved in methanogenesis and corresponding genes detected in the metagenomes of NAPDC, SCADC, and TOLDC. **(A)** Pathways for acetotrophic, hydrogenotrophic, and methylotrophic methanogenesis. **(B)** Heatmap showing the presence of genes and putative genes detected in dominant members of the methanogenic enrichment culture metagenomes. The colour scale indicates the inferred enzyme completeness in the case of enzymes with multiple subunits. Mc; *Methanoculleus*, Ms; *Methanosaeta*, Mr; Cand. *Methanoregula*, Oth; other methanogens. Enzyme abbreviations: Fmd/Fwd, Formylmethanofuran dehydrogenase; Ftr, Formylmethanofuran formyltransferase; Mch, Methenyl-H_4MPT cyclohydrolase; Mtd, F420-dependent methylene-H_4MPT dehydrogenase; Hmd, H_2-dependent methylene-H_4MPT dehydrogenase; Mer, Methylene-H_4MPT reductase; Mtr, methyl-H_4MPT methyltransferase; Mcr, Methyl-coenzyme M methylreductase; CODH/ACS, Carbon-monoxide dehydrogenase/Acetyl-CoA synthase; AK, Acetate kinase; Pta, Phosphotransacetylase; Mta, Methanol:coenzyme M methyltransferase; Hdr, Heterodisulfide reductase. AK was omitted from the heatmap as AK genes identified were found not to affiliate with methanogens and are thus not representative of this pathway.



Supplementary Figure S5. Relative abundance in NAPDC, SCADC, and TOLDC of groups of genes involved in key metabolic pathways, predicted using the KEGG pathway module (Kanehisa *et al.* 2012).

Supplementary References

Davidova IA, Gieg LM, Nanny M, Kropp KG, Suflita JM. (2005). Stable isotopic studies of *n*-alkane metabolism by a sulfate-reducing bacterial enrichment culture. *Appl Environ Microbiol* **71:** 8174-8182.

Heider J, Fuchs G. (1997). Microbial anaerobic aromatic metabolism. *Anaerobe* **3:** 1-22.

Holowenko FM, MacKinnon MD, Fedorak PM. (2000). Methanogens and sulfate-reducing bacteria in oil sands fine tailings waste. *Can J Microbiol* **46**: 927- 937.

Hug LA, Beiko RG, Rowe AR, Richardson RE, Edwards EA. (2012). Comparative metagenomics of three Dehalococcoides-containing enrichment cultures: the role of the non-dechlorinating community. *BMC Genomics* **13:** 327.

Kanehisa M, Goto S, Sato Y, Furumichi M, Tanabe M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res* **40:** D109-D114.

Leutwein C, Heider J. (1999). Anaerobic toluene-catabolic pathway in denitrifying *Thauera aromatica*: activation and beta-oxidation of the first intermediate, (R)-(+)-benzylsuccinate. *Microbiology-Sgm* **145:** 3265-3271.

Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M *et al.* (2008) The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9:** 386.

Siddique T, Fedorak PM, MacKinnon MD, Foght JM. (2007). Metabolism of BTEX and naphtha compounds to methane in oil sands tailings. *Environ Sci Technol* **41:** 2350-2356.

Tan B, Dong X, Sensen CW, Foght J. (2013). Metagenomic analysis of an anaerobic alkane-degrading microbial culture: potential hydrocarbon-activating pathways and inferred roles of community members. *Genome* **56:** 599-611.

Taubert M, Vogt C, Wubet T, Kleinsteuber S, Tarkka MT, Harms H *et al.* (2012). Protein-SIP enables time-resolved analysis of the carbon flux in a sulfate-reducing, benzene-degrading microbial consortium. *ISME J* **6:** 2291-2301.

Widdel F, Bak F. (1992). Gram-negative mesophilic sulfate-reducing bacteria. In: Balows A, Trüper HG, Dworkin M, Harder W, Schleifer KH (eds). *The Prokaryotes*. Springer-Verlag: New York, USA. pp. 3352-3378.

Winderl C, Penning H, von Netzer F, Meckenstock RU, Lueders T. (2010). DNA-SIP identifies sulfate-reducing Clostridia as important toluene degraders in tar-oil-contaminated aquifer sediment. *ISME J* **4:** 1314-1325.