We thank the thoughtful suggestions and constructive comments, which help to improve our manuscript. We also appreciate the insights of the reviewer.

We have followed the detailed suggestions, and addressed all the points that were raised, and have rectified all the mistakes and errors through this revision. Hopefully, this can serve as a response to answer for all the major points.

Anonymous Referee #1

Review of manuscript bg-2018-470 Core bacterial microbiome of petroleum reservoirs in China.

The manuscript contains interesting data and is sufficiently well written such that it should be suitable for publication after minor modifications. The main contributions of the manuscript are to help define a common core microbial community in petroleum reservoirs and to hypothesize that the dominant methanogenic biochemical pathway and associated microorganisms shift in water versus oil phase microorganisms, and that the relative importance of hydrogenotrophic and acetoclastic pathways is different in low versus high temperature reservoirs.

Major comments:

Nowhere in the text are the temperatures that correspond to low, moderate, and high petroleum reservoirs described. The revised manuscript should specify temperatures in the abstract and the introduction as a minimum.

Reply: Thanks to this suggestion. We added the detail temperature information in the Abstract (Line 23) and Introduction (Line 101) according to the reviewer's suggestion.

Since aqueous phase and oil phase microbes were recovered from only 7 locations consisting of 2 low, 4 moderate, and one high temperature reservoirs it is premature to state conclusions that the relative importance of hydrogenotrophic and acetoclastic pathways is different in low versus high temperature reservoirs. Rather, the revised manuscript should discuss the observations made by inspecting these data and offer a hypothesis that requires further testing.

Reply: Thanks for the comments here. Since the limited locations and samples in this study, it is truly premature to make a conclusion as the reviewer pointed out. As a result, we summarized the observations and confirmed the distribution pattern by three taxonomical investigating methods, e.g., *mcrA* gene-based (clone library), methanogen 16S rRNA gene-based (clone library), and MiSeq-based archaea 16S rRNA gene methods (Fig. 3). All of these three methods support a similar community structure, suggesting that the results in our manuscript are reliable and could serve as a hypothesis that requires further testing with more samples for the ubiquity of the observation.

Table 1 shows that the concentration of acetic acid in the water phase of these 7 samples does not correlate with the reservoir temperature and does not support the claims made regarding the importance of acetoclastic methanogenesis derived from the genetic data. This should be discussed.

Reply: Thanks for the insightful comments. We double-checked the original acetate data, and are sure about the reliability of them. The chemical datasets were measured from the aqueous phase of each well production water sample by ion chromatography as stated in the Materials and Methods. We observed the community shifting from hydrogenotrophic methanogenesis to acetoclastic methanogenesis in high temperature reservoirs from oil phase to water phase, but this is not necessarily equivalent to that acetoclastic methanogenesis is more important in high temperature reservoirs and the corresponding acetate-utilizing rate will be faster than the other eco-niches. Meanwhile, as in a complex microbial community, the other members and processes could also potentially contribute to the consumption of acetate, such as syntrophic acetate oxidization pathway coupling with hydrogenotrophic methanogenesis. Therefore, it is not suitable for us to make a reasonable conclusive deduction at this moment.

Without the *in situ* microbial activity data and other quantitative measuring methods on the chemical reactions (stable isotope labeling), it is still unreliable to address the direct connection between the microbial function/activity and acetate concentration in any petroleum reservoirs. In the manuscript (Line 330-331), we introduced more discussion about the acetoclastic methanogenesis and syntrophic acetate oxidization processes.

Another issue is that the production of glycine betaine as an osmoprotectant in high salinity petroleum reservoirs has been shown to serve as the main source of methane,
via methylotrophic methanogenesis, in several reservoirs. The reservoirs studied here are low salinity and no evidence of methylotrophic methanogenesis was found, but the
discussion of a core microbial community should be improved by discussing genetic data from higher salinity reservoirs as it compares with the results of this study.

Reply: Thanks for the comments. Actually, we found the molecular evidence of methylotrophic methanogenesis in P1 aqueous and oil phase samples; they included different groups within Methanosarcinales (Fig. 3). We could not conclude that they are obligate methylotrophs, while, to some extent, the existence and abundance in certain petroleum sites suggest the availability of methyl-containing compounds, e.g., glycine betaine in saline environments (Ollivier et al., 2010, oil reservoir ecosystem). It is suggested that this osmoprotectant could be accumulated or synthesized by the halophiles, and the degraded methyl-compounds from fermentative processes will subsequently fuel the growth of the methylotrophic methanogens. We added the corresponding discussion in the manuscript (Lines 307-316).

Are there any prior publications describing the recovery of microbial cells from oil-phase samples, or is this a new technique reported for the first time? If this technique is new, that should be highlighted in the revised manuscript.

Reply: Thanks for the comments. To the best of our knowledge, there is not any report available on obtaining microbial cells from oil-phase samples directly. The reason might be: 1) difficulties to preserve production fluids (oil/water mixture) and maintain the viability of the cells when retrieving them from high pressure/temperature oil-wells; 2) good fractionation method to separate the cells from oil phases through extraction. The current studies on oil-phase microbes are all based on metagenomics approach.

Line 86 change pervious to previous.
Reply: Revised as suggested (Line 87).

Line 275 replace ture with true.
Reply: Corrected as suggested (Line 286).

Anonymous Referee #2

In this manuscript, Zhichao Zhou et al describe the geochemistry and microbial community diversity of several petroleum reservoirs characterized by a broad range of temperatures. For microbial community analyses, the aqueous and oil phases were separated and treated independently. Analyses were essentially based on 16S rRNA and *mcrA* gene sequencing. In addition, qPCR for the *mcrA* gene were done, although the results of this analysis do not appear to be extensively discussed in the main text. A major conclusion of the work is that despite the distinct geographical locations and different physical-chemical parameters, the analyzed reservoirs shared a common microbiome represented by a small number of OTUs of high relative abundance. Shifts of the methanogenesis mechanisms between the aqueous and oil phase were observed in low and high temperature reservoirs (but not in those of moderate temperature).

Major concerns:

Overall, the findings are interesting, but there are, however, several aspects that need more careful consideration. The Introductions makes the case for MEOR: however, in the study rationale (line 95-98) there is not apparent connection to MEOR, and this is also not mentioned again in the conclusions. Is there an impact of the findings of this work on MEOR? Could the observed similarities between the microbial communities of the reservoirs be actually caused by the many years of enhanced oil recovery applied at those sites?

Reply: Thanks for the comments. These reservoirs have been under enhanced oil recovery (EOR) practice for many years, such as flooding with chemical and water. We could not rule out the possibility completely in that the long-time EOR could result in microbial community similarity observed in this study. There is a reasonable possibility in that substantial portions of aerobic bacteria discovered in the core bacteria microbiome across different petroleum reserviors could be introduced by the exogenous sources through the water flushing. We added this into the Discussions (Lines 263-266). Meanwhile, as indicated by the reviewer, there are not so many interpretations on the connection between MEOR and the microbial

community and function investigated in this study. We also added some information into the Discussions on this in the manuscript (Lines 97-99, 257-262).

Overall, the manuscript needs some revision with respect to writing. For some sentences, it is not right away clear if they refer to previous papers, or to results of the present study (for ex. line 30-32, 271-272).

Reply: Thanks for the suggestions. We added "previous reports" to the sentences to make it clear (Lines 29, 278)

It was demonstrated in many other studies that the PCR reaction is introducing biases: clone libraries and MiSeq data should be therefore interpreted as 'relative abundance'. With the exception of methanogenesis, which benefits from *mcrA* analyses, all other function assignments are done based on short 16S rRNA gene sequences which give taxonomic resolution at phylum level. This should be revised, and sequence data should be more cautiously interpreted. For example, there is no solid basis to assign all those detected phyla as hydrocarbon degraders (line 28-30).

Reply: Thanks for the comments. We used the 515F/909R based-MiSeq sequencing to generate short 16S rRNA gene reads (which is the longest read length that could be achieved by High-throughput Sequencing method currently, see http://www.earthmicrobiome.org/protocols-and-standards/16s/. The others would be even shorter, such as 300bp or 200 bp). We could obtain the taxonomic assignments down to the level of genus for certain groups by using QIIME and the SILVA database (the most comprehensive 16S rRNA database better than others, such as Greengenes and etc). We did not simply assign the function to the microbes at the phylum level but we picked the most abundant genera/core OTU, and found their functional roles according to the reported cultured strains (Tables S5 and S6) (We will also have a review paper on this in the near future). Nevertheless, it is not allowed to assign any functions to certain microbial groups simply based on the closest cultured strains in the same genus/family/order; at the current stage, we think that this is the most careful or reasonable way to deal with the "digital datasets".

It is stated that methanogenesis is a dominant process (Abstract), but there are not too many arguments provided for this; MiSeq analysis show a high relative abundance of Proteobacteria in all samples (Fig. 1), and these are not methanogens. Also, some samples contain very high amounts of sulfate, comparable to sea water (P1 and P5); in the same samples, nitrate is not detectable. If one assumes oxygen is also absent, conditions are favorable for sulfate reduction rather than for methanogenesis. This is not discussed in the manuscript.

Reply: Thanks for the comments. The archaeal methanogenesis is not the most dominant process, and their relative abundance could not compete with bacterial ones. We rephrase the sentence here, that, **the methanogenesis process shifts from the dominant**

hydrogenotrophic pathway in aqueous phase to acetoclastic pathway in oil phase in high-temperature reservoirs (Line 32, 284).

In the low oxygen subsurface petroleum environments, as the reviewer suggested, high sulfate concentrations could fuel the growth and activity of sulfate-reducing bacteria, and this will compete with methanogenesis process. We do find sulfate reducer increase in these samples, e.g., a large proportion of Firmicutes and Deferribacteres in the P5A for the sulfate reduction. We added this part of discussion to the manuscript (Line 300-307).

Currently the reader cannot appreciate if the method used to separate aqueous and oil phases (line 109-110) may alter the community structure. More details should be given, especially the times employed. Was this method tested to make sure cells are not lysed during heating? For how long was the mixture kept at 50C? What exactly are the 'undetermined results' (line 210-211) that were omitted?

Reply: Thanks for the comments. In the actual operations, there is not a defined time for the heating. The purpose of this step is to heat the cool and solidified petroleum fluids into semi-fluids, and to benefit for the downstream fractionation. We added the details into the manuscript (Line 114-117).

Normally, we did three replicates for one qPCR to determine the gene quantities. However, for certain experiments, there will be one data that is significantly deviated from others, and we deleted these data from the replicate groups. The undetermined result means a result that is under the detection limit. We added the interpretation in the manuscript (Line 218).

'Core microbiome' is used often in the text: this should be defined. Does it refer to all taxa with a relative abundance over 0.1%? What does 'quantity requirement for quality control' means exactly? Other inconsistencies: in the Abstract, enrichment cultures are mentioned (line 29), but there is nothing mentioned about enrichments in the rest of the manuscript; Salmonella and tuberculosis are associated with archaeal communities (line 355 – 359)? The equilibrate claim at line 304-305 is at odds with the conclusion at line 371-372. Figure 3, header is missing for the last column?

Reply: The core microbiome actually indicates the OTUs that are shared among all petroleum components from all the samples in this study. The "quantity requirement for quality control" means that we pool the DNA extraction from several repetitive experiments to meet the quantity requirement and then use the DNA to conduct quality control (Line 128).

In the abstract, the enrichment cultures are referring to the previously reported studies (We added this information accordingly, see Line 29). *Salmonella* and tuberculosis are usually not

associated with archaeal community; however, this is the interpretation result from the Tax4Fun and LEfSe. Since FTU values of archaeal communities (fraction of taxa that could be mapped to existing KEGG pathway) were unevenly distributed from sample to sample, the reliability of these functional predictions is in question (Line 393-395). So, I didn't go much into the result discussion for this part, and also state the shortage of this analysis which is based on the current collected database that is far from completeness (Line 395-397).

In this study, we observed the shift of dominant methanogenesis pathway between aqueous and oil phases within samples by temperature, but we do not know whether the temperature is directly/indirectly involved with the shift and influence mechanisms. This needs further observations and studies. We change the expression of related contents (Line 333-334).

The last column is for sample P7A and P8O, which are from different petroleum reservoirs. Fig 3 reflects the compositional shift of methanogenic archaeal community of aqueous/oil phases of individual samples. So, we did not include the last two separated samples which lacked the corresponding aqueous/oil phase samples.

Identify the core bacterial microbiome of hydrocarbon degradation and a shift of dominant methanogenesis pathways in oil and aqueous phases of petroleum reservoirs of different temperatures from China

Zhichao Zhou[1], Bo Liang[2], Li-Ying Wang[2], Jin-Feng Liu[2], Bo-Zhong Mu[2], Hojae Shim[3], and Ji-Dong Gu[1,*]

[1] Laboratory of Environmental Microbiology and Toxicology, School of Biological Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong SAR, Hong Kong, People's Republic of China

[2] State Key Laboratory of Bioreactor Engineering and Institute of Applied Chemistry, East China University of Science and Technology, Shanghai 200237, People's Republic of China

[3] Faculty of Science and Technology, University o Macau, Macau, People's Republic of China

*Correspondence to: Ji-Dong Gu (jdgu@hku.hk)

Abstract

Microorganisms in petroleum reservoirs play significant roles in hydrocarbon degradation and the terminal electron accepting process of methanogenesis, they also contribute to microbial enhanced oil recovery (MEOR) worldwide with great economic and environmental benefits. Here, a molecular investigation, using the 16S rRNA and *mcrA* gene profiles based on MiSeq sequencing and clone library construction methods, was conducted on oil and water (aqueous) phases of samples of high (82-88°C), moderate (45-63°C) and low temperatures (21-32°C) from seven petroleum reservoirs in China. A core bacterial microbiome with a small proportion of shared operational taxonomic unit (OTU), but a high proportion of sequences among all reservoirs was discovered, including aerobic degraders, sulfate/nitrate-reducing bacteria, fermentative bacteria and sulfur-oxidizing bacteria distributed mainly in *Proteobacteria*, *Bacteroidetes*, *Deferribacteres*, *Deinococcus-Thermus*, *Firmicutes*, *Spirochaetae*, and *Thermotogae*. Their prevalence in the previously reported petroleum reservoirs and successive enrichment cultures suggests their common roles and functions involved in aliphatic and aromatic hydrocarbon degradation. The methanogenic process generally shifts from the dominant hydrogenotrophic pathway in aqueous phase to acetoclastic pathway in oil phase in high-temperature reservoirs, but the opposite was true for low-temperature reservoirs. No difference was detected between two phases in moderate temperature reservoirs. Physicochemical factors, including pH, temperature, phase conditions, and nitrate, Mn^{2+}, and Mg^{2+} concentrations are the main ones correlated to the microbial compositional and functional profiles significantly. LEfSe analysis shows distribution differences of microbial groups towards pH, temperature, and oil/aqueous phases. Tax4Fun functional profiling indicates major functional metabolism differences between the two phases, including amino acids, hydrocarbons in oil phase, and carbohydrates in aqueous phase.

Keywords: Reservoir; Core microbiome; Oil and aqueous phases; Temperature; Methanogenesis; Hydrocarbon

1 Introduction

Microbial enhanced oil recovery (MEOR) is one of the most feasible and profitable technologies for extracting residual oil from low-productivity reservoirs by utilizing microorganisms for hydrocarbon transformation via fermentation and methanogenesis as the major functional players in MEOR applications (Magot et al., 2000; Mbadinga et al., 2011; Mesle et al., 2013). Implementation of MEOR could also attenuate side-effects by injection of nitrate into oil wells to stimulate the propagation of nitrate-reducing bacteria (NRB) and inhibit the growth of sulfate-reducing bacteria (SRB), subsequently reducing sulfidogenic negative effects on corrosion and oil quality deterioration (Gao et al., 2013). Currently, investigations on compositional and functional profiles of petroleum reservoir microbiome have attracted attention from both scientists and engineers because the information will improve the current understanding on microbial diversity and function in oil reservoir systems and implement MEOR through manipulating microbial community composition and activities (Magot et al., 2000; Mesle et al., 2013; Youssef et al., 2009).

Petroleum reservoirs are often characterized as high temperature, and high pressure subterranean ecosystems, but low-temperature reservoirs also exist around the world (Grabowski et al., 2005; Li et al., 2013; Voordouw et al., 1996; Li et al., 2014). The physicochemical and geographic differences among petroleum reservoirs together with exogenous injection of microorganisms and nutrients during MEOR operation could affect the composition and functional groups of the indigenous microbial community. Many previously discovered bacterial and archaeal groups have important functional roles, e.g.,

sulfate-reducing, nitrate/nitrite-reducing and fermentative bacteria, thermophilic archaeal hydrocarbon degraders and methanogens as well as exogenous aerobic hydrocarbon degraders in the microbial community (Gao et al., 2015a; Gao et al., 2015b; Lenchi et al., 2013; Li et al., 2012a; Orphan et al., 2003; Wang et al., 2012; Grabowski et al., 2005; Pham et al., 2009). The common and specific roles and functions of these microorganisms in petroleum reservoirs are of great value to be comprehensively examined.

Methanogens cooperate with hydrocarbon degraders and fermentative bacteria synergistically to make the overall degradation processes thermodynamically favorable under largely anaerobic conditions (Mbadinga et al., 2011). Fermentative products from hydrocarbon degradation provide the essential growth substrates for methanogens with two major pathways, namely hydrogenotrophic and acetoclastic methanogenesis, and in many cases the syntrophic acetate oxidization pathway coupling with hydrogenotrophic methanogenesis is prevalent in both *in situ* and enrichment incubation *ex situ* (Mesle et al., 2013; Liang et al., 2015; Wang et al., 2012; Lee et al., 2015; Mayumi et al., 2011). However, very little is known on the influencing factors that shape the methanogenesis pathways and methanogen community structure and composition.

Petroleum reservoir is a complex system consisting of porous sandstones with oil, water and air. The microorganisms attached to the oil phase of petroleum fluids are largely neglected in most of the previous investigations due to technical difficulties in DNA extraction and sequencing (Kryachko et al. 2012). Oil-attached microorganisms influence the oil-water interface properties by the production of biosurfactants and metabolites or forming adsorbed colloidal particles to enhance oil recovery performance (Kobayashi et al., 2012; Kryachko et al., 2012; Wang et al., 2014). The distribution, function, and contribution of these microorganisms to MEOR success could be very different from those of the aqueous phase

(Kobayashi et al. 2012). Oil-degrading microorganisms, including thermophilic hydrocarbon-degrading archaea and methanogens, play different roles in the MEOR process (Mbadinga et al., 2011). Their compositional patterns and functional profiles towards temperature and oil/aqueous phase, are of great value to understand the mechanism of MEOR operation. Knowledges on the major microbial drivers, their potential functions, distribution characteristics and changing patterns towards environmental parameters should be one of the research directions for better understanding the MEOR process. In this study, the objectives were to study the compositional and functional differences of petroleum reservoir inhabiting microorganisms under different temperatures (high: 82-88°C, moderate: 45-63°C and low: 21-32°C); the methanogenesis pathways in oil and aqueous phases of these samples; and the influence of physicochemical factors on microbial community composition.

2 Materials and Methods

2.1 Characterization of geographic properties of sampling reservoirs

Petroleum production fluid samples were collected from seven areas across China, covering oilfields of different geographical locations and temperatures. The reservoirs and crude oil properties together with the aqueous phase chemical characteristics of this study are listed in Table 1. Detail reservoir properties are described in the Supplementary Materials.

2.2 Sample collection, ion concentration measuring, and DNA extraction

Each sample containing a mixture of crude oil and water was collected on site into a sterilized container after flushing each wellhead for at least 3-5 min. The containers were screwcapped to avoid air invasion and transported to the nearby laboratory immediately for further processing. The mixture was gently heated to 50°C to make it semi-fluid and then separated into oil and aqueous phases in a separatory funnel. The heating was operated as short as

possible according to the solidification degree of individual petroleum fluid samples, in order to reduce the lysis effect of microbial cells within. Ion concentrations of the aqueous phase were measured using Dionex 600 ion chromatography (Triad Scientific, Inc., Manasquan, NJ, USA) following the manufacturer's instructions.

To obtain aqueous phase DNA, the aqueous phase of a sample after oil and water separation was first filtered through a 0.22 polycarbonate membrane filter and a slice of the membrane was used to extract DNA by AxyPrep™ Bacterial Genomic DNA Miniprep Kit according to manufacturer's instruction (Axygen Biosciences, USA). For the oil phase, three volumes of isooctane (2, 2, 4-trimethylpentane) were used to dissolve the crude oil, and then centrifuged at 5000 rpm for 30 min to concentrate non-dissolved particulate and microbial cells for at least three times to obtain enough materials. Repetitive DNA extractions were then conducted on these materials and combined to meet the quantity requirement for the downstream quality control. Finally, all DNA samples from aqueous and oil phases were measured by Nanodrop for concentration and checked for DNA integrity by electrophoresis.

2.3 Clone library construction and analysis

The methyl-CoA enzyme-encoding gene (*mcrA*) primer pair, ME3MF&ME3MF-e/ME2r'(Narihiro and Sekiguchi, 2011) and the 16S rRNA gene primer pair targeting methanogens, MetA86f/ARC915r (Wright and Pimm, 2003; Narihiro and Sekiguchi, 2011; Yu et al., 2008), were used to generate gene clone libraries for each sample (aqueous and oil phases separately) according to the method established previously (Zhou et al., 2015) (Table 2).

FunGene pipeline chimera check was applied to check *mcrA* gene sequences using the UCHIME *de novo* mode (Edgar et al., 2011). USEARCH software was used to check

chimeras of methanogenic 16S rRNA gene sequences using QIIME compatible SILVA 119 release SSURef database ("rdp_gold fasta") file as a reference. Then, for *mcrA* gene sequences, *de novo* OTU picking method was applied by QIIME at a cutoff value of 0.05 (Caporaso et al., 2010). Representative OTU sequences were aligned and inserted into the *mcrA* gene ARB database by maximum parsimony method without changing the initial tree topology (Angel et al., 2012; Ludwig et al., 2004). The phylogenetic affiliation was assigned, and taxonomic composition results were processed by QIIME accordingly (Figure S1). For methanogenic 16S rRNA gene sequence clustering and diversity analysis, the same method was applied via QIIME as described in the followings.

2.4 MiSeq sequencing and QIIME based analysis

The prokaryotic universal primer pair 515F/909R (Caporaso et al., 2012; Wang and Qian, 2009) and archaeal universal primer pair Arch347F/Arch806R (Takai and Horikoshi, 2000) were used to amplify samples of this study (both with barcodes attached to the forward primers, Table 2). Two PCR reactions were conducted and then combined to yield enough PCR products to compromise variations between different batches. Then, pooled PCR products of each sample with approximately 100 ng or 200 ng DNA were subjected to one MiSeq run. Sequencing samples were prepared using TruSeq DNA Kit according to the manufacturer's instruction. The library was uploaded to an Illumina MiSeq platform for sequencing with reagent kit v2 (2× 250 bp) or v3 (2× 300 bp) as described in the manufacturer's manual.

After merging paire-end reads from raw sequencing data by FLASH-1.2.8, fastx-toolkit was applied to split merged reads from one run into individual samples according to the primer barcodes (Magoc and Salzberg, 2011). Then, all sequences were split into each library with the name of each sample attached according to the barcode map using QIIME command

"split_libraries" (Caporaso et al., 2010). The criterion for filtering out underqualified sequences was "-s 15 -k -a 6 -r -l 150 -b 12 -M 5 -e 0". Chimera checking was conducted by USEARCH software using QIIME compatible SILVA 119 release SSURef database ("rdp_gold" fasta) file as the reference (Edgar, 2010). Clustering, picking OTU, taxonomy assignment, aligning, filtering alignments and phylogenetic tree construction, taxonomic composition summarizing, alpha and beta diversity analyses were conducted step-by-step by the QIIME pipeline with QIIME compatible SILVA 119 SSURef database as the reference (Caporaso et al., 2010). In clustering, "pick_open_reference_otus.py" command was used to conduct OTU dividing and BLAST method was used to assign taxonomy to input sequences. This subsampled open reference OTU picking method was the performed-optimized and optimal strategy suggested by the developers (Rideout et al., 2014). After OTU table "biom" files were generated, in order to get the bacterial community composition information from prokaryotic 16S rRNA gene primer amplified libraries, "filter_taxa_from_otu_table.py" command was introduced to only retain bacterial OTUs in the "biom" file. Similarly, exclusive archaeal and methanogenic OTU table files could also be processed from archaeal 16S rRNA gene primer amplified libraries. Summary information of each sample OTU abundance could be calculated by "biom summarize-table" command, and then the lowest number among all samples was chosen as the subsampling size to make each library acquire an even size using "multiple_rarefaction_even_depth.py" command. The taxonomic compositional table was drawn according to the subsampled biom file. Since there was no lanemask file available in this SILVA compatible 119 release SSURef database, alignment filtering method was performed independently with entropy threshold as 0.1 and gap filter threshold as 0.9 after obtaining aligned sequences by PyNAST method. Diversity parameters of each library could be generated by alpha diversity calling, with rarefaction curve, Good's coverage value, Shannon, Chao1, Simpson and PD whole tree indices calculated. Beta diversity, which delineates the dissimilarity relationship among samples, was generated and

visualized through unweighted & weighted UniFrac matrix and non-phylogenetic Bray-Curtis matrix method. The pairwise-shared OTU numbers were calculated from "biom" by the command "shared_phylotypes.py". Core microbiome (shared OTU table in a specific sample category) was identified by "compute_core_microbiome.py" command.

2.5 Diversity and statistical analysis

The statistical significance of community composition of samples among different categories was valued by anosim and adonis method, implemented in "compare_categories" command in QIIME. Mantel Test was used to compare the distance matrix of physicochemical parameters and UniFrac/Bray-Cutis distance matrix from beta diversity analysis by QIIME. For aqueous phase samples, both of the *in situ* physicochemical parameters and ion concentrations were used in the analysis; while only *in situ* physicochemical parameters were included for oil phase samples. The compositional bar chart and bubble chart were modified and illustrated from taxonomic summary results. Tax4Fun was used to predict the functional capabilities based on abundance profiles of microbial 16S rRNA gene datasets (Asshauer et al., 2015). LDA Effect Size (LEfSe) analysis was applied to illustrate the biomarker species with high statistical significance in different sample categories and the functional profiles statistically distributed in different sample categories (Segata et al., 2011).

2.6 Quantitative PCR on *mcrA* gene abundance

The quantitative PCR measurement was conducted using iTaq™ Universal SYBR® Green Supermix Kit (BIO-RAD). The qPCR mixture contained in 15 μl: 7.5 μl of supermix, 16 μg of BSA (Roche), and 1 μM final concentration of primer pair [ME3MF&ME3MF-e (250:1) and ME2r']. Annealing temperature was set to be the same as the clone library PCR setting, and the rest thermocycling settings were according to the manufacturer's instructions. Randomly picked one pMD18-T plasmid with *mcrA* gene inserted was used to make the

standard curve. The DNA template concentration was adjusted to 0-40 ng/μl. Results, which deviated significantly from values in the replicate groups, were omitted and undetermined results (under the detection limit) were also deleted. The property of final adjusted standard curve is $r^2 = 0.995$ and Eff% = 83.32.

2.7 Sequencing result deposition

For clone library sequences, KT314862-KT315353, were assigned to methanogenic 16S rRNA gene sequences; KT314340-KT314835 were assigned to *mcrA* gene sequences. For high throughput sequencing data based on Illumina MiSeq platform, PRJEB9855 was assigned to prokaryotic 16S rRNA gene sequencing data; PRJEB10996 was assigned to archaeal 16S rRNA gene sequencing data.

3 Results and Discussions

3.1 Common OTU among different categories and core bacterial microbiome

Community composition results showed that 21 bacterial phyla were obtained with an average abundance of more than 0.1% (Figure 1a), and three major archaeal phyla were *Thaumarchaeota*, *Euryarchaeota*, and *Crenarchaeota* (Figure 1b). Pairwise-shared OTU numbers of all samples indicated that, irrespective of combinations between aqueous and oil phase samples, the average numbers of shared bacterial OTUs ranged from 199.9 to 292.4, accounting for 26.6-36.2% of the total OTU numbers within individual samples; average numbers of shared archaeal OTUs ranged from 1.8 to 11.9, accounting for 8.5-23.4% of total OTU numbers within individual samples (Table S4). Core bacterial OTU numbers among aqueous, oil and all samples were 73, 57 and 46, which accounted for 7.1-10.1%, 5.9-10.3% and 4.5-8.3% of OTU numbers in individual samples. Core archaeal OTU number among

aqueous samples was only 3, accounting for 3.9-8.1% of OTU numbers in individual samples, and no archaeal OTU was shared among oil phase samples.

However, by investigating taxonomic profiles of core bacterial OTUs, the shared OTUs were 49, 41 and 34 genera in aqueous, oil and all samples, corresponding to 65.5%, 59.9% and 58.8% of average sequences in the total bacterial community, respectively (Tables S5 and S6). Most of the core bacterial OTUs belonged to the most abundant 36 genera, of which numbers of shared genera among aqueous, oil and all samples were 28, 23 and 23, respectively (Figure 2, and Tables S5 and S6).

There was no significant difference of shared bacterial OTU numbers within and between aqueous and oil phase samples, suggesting a core microbiome was shared among all components. The core OTUs covered around 2/3 of the total bacterial sequences, even though the percentages of core/total OTU number for individual samples were 4.5% to 10.3%. The core microbiome shared among all petroleum reservoirs could be the key participants mediating critical microbial processes, such as activation, degradation, fermentation, oil-emulsification and methane generation (Yamane et al., 2008; Wang et al., 2014; Pham et al., 2009; Orphan et al., 2000; Magot et al., 2000). This spectrum of core microbiome shares common functional roles in facilitating MEOR and is modified by the *in situ* physicochemical conditions of different reservoirs (Figure 2 and Table S5). It is important to connect the major microbial players, including their community compositions and specific functional capacities, to the interpretation of MEOR processes in the petroleum reservoirs. Meanwhile, the core microbiome serves as a good study target for primarily simplifying microbial participants and their roles in the petroleum reservoirs, and is useful for modeling and monitoring the MEOR processes for petroleum reservoirs from different locations. Moreover, substantial portions of aerobic bacteria being discovered in the core bacterial

microbiome across reservoirs imply that exogenous bacteria introduced into reservoirs by water flushing can be also represented in the core composition and play important roles in enhancing oil recovery (Gao et al., 2015b; Youssef et al., 2009).

Bacterial community distribution in aqueous phase showed correlation with temperature by Mantel Test using unweighted UniFrac matrix method, and temperature also significantly affected the dissimilar distance matrix of PCoA coordinates of the bacterial community by both adnois and anosim statistical analyses (Tables S3, S7, and S8). However, temperature changed the proportions of the taxa of the core microbiome significantly, but not the presence/absence (Tables S3, S7, and S8). Consequently, the common genera discovered from this study show that despite geographic and physicochemical differences, there is a core microbiome with small OTU numbers but large sequence proportions in these petroleum reservoirs across China which are possibly responsible for mediating hydrocarbon degradation processes. Prevalent core genera discovered in this study also appear to be the most abundant components as previously reported in petroleum reservoirs within and beyond China (Gao et al., 2015a; Gao et al., 2015b; Li et al., 2012a; Orphan et al., 2003; Orphan et al., 2000; Tang et al., 2012; Yamane et al., 2008; Dahle et al., 2008; Li et al., 2015; Liang et al., 2015; Li et al., 2014) and enrichment cultures from petroleum reservoirs worldwide (Gray et al., 2011; Li et al., 2012b; Liang et al., 2015; Wang et al., 2011).

3.2 Shift of major methanogenesis pathways between oil and aqueous phases

The methanogenic process generally shifts from the dominant hydrogenotrophic pathway in aqueous phase to acetoclastic pathway in oil phase in high-temperature reservoirs. The opposite is true for low-temperature samples, but no difference was detected between the two phases in moderate temperature reservoirs (Figure 3). The different patterns of shifting of the major methanogenesis mechanisms in aqueous and oil phases were evident in different

temperature petroleum reservoirs of this study. Quantitative measurements of *mcrA* gene in both aqueous and oil phases of all samples are summarized in Supplementary Material.

The underlying methanogenesis mechanism could change substantially in reservoirs with distinctive geochemical properties (Magot et al., 2000). The dominance of methanogenic process mediated by methylotrophic methanogens is rarely observed in petroleum reservoirs. It is claimed that there is a very low possibility that methyl-compounds could be generated during the degradation of kerogen (Mesle et al., 2013). However, the dominance of obligate methylotrophic *Methanolobus* was found in all water, rock and coal samples in a coalbed methane reservoir within Eastern Ordos Basin of China (Guo et al., 2012). It could be deduced that the prevalence of methylotrophic methanogenesis under certain conditions is directly fueled by the existence and availability of methyl-containing substances. Since that bacterial relative abundances in petroleum reservoir samples are always higher than those of methanogenic archaeal community, the methanogenesis process might not be the dominant process among all microbial processes. For instance, in P1 and P5 aqueous samples, sulfate concentrations are considerably high (Table 1), geochemical conditions are more favorable for sulfate reduction than methanogenesis. A large proportion of Firmicutes and Deferribacteres in P5A are potentially responsible for the activities of sulfate reduction according to their relative fractions (Figure 1) and functional capacities (Figure 2). It is suggested the methylated compounds could be produced by the degradation of glycine betaine (an osmoprotective agent), that are accumulated or generated by halophiles in saline petroleum (Ollivier and Alazard, 2010). Subsequently, it fuels the growth of methylotrophic methanogens in certain petroleum environments. Although, no molecular or chemical evidence was reported for this process in petroleum environments, a pioneer study on hydraulic fracturing in shales has detected glycine betaine as the major known osmoprotectant in the produced fluids, and the combined metagenomic and metabolite

findings suggests the similar glycine betaine utilization pathways in fractured shales (Daly et al., 2016).

Numerous studies *in situ* or in enrichment incubation microcosms on aqueous phase of reservoir fluids claimed that syntrophic acetate oxidation associated with hydrogenotrophic methanogenesis is the major hydrocarbon degradation pathway (Wang et al., 2012; Mayumi et al., 2011; Lee et al., 2015; Mbadinga et al., 2012; Gray et al., 2011). The dominance of hydrogenotrophic methanogenesis in subsurface ecosystem could result from the external hydrogen originated from maturation of organic matters and/or mineral hydrolysis (Head et al., 2003), and the synergistic effect in association with acetate oxidizers, whereby acetate was firstly oxidized to H_2 and CO_2, then utilized by methanogenesis (Liu and Whitman, 2008). Additionally, a stable isotope labeling experiment on oil-degrading microcosm showed that despite coexistence of acetoclastic methanogenesis and acetate syntrophic oxidization in the initial state, the latter process prevailed over the former one when introducing low initial acetate concentration (Gray et al., 2011). The above evidence suggests that acetate syntrophic oxidization could exceed acetoclastic methanogenesis and contribute substrates H_2 and CO_2 to potentially favor hydrogenotrophic methanogenesis process.

It is still difficult to determine whether the temperature has directly or indirectly been involved in the alteration of the methanogenesis pathways. Reservoir fluid constituents may affect methanogenic degradation because crude oil and creosote inhibit acetoclastic methanogenesis (Warren et al., 2004) and volatile hydrocarbons (nC$_5$-nC$_{10}$) inhibit methanogenic degradation rate without changing the abundances of both hydrogenotrophic and acetoclastic methanogens (Sherry et al., 2014). Since most currently available community studies are based on the microbial assemblages within injection or production water, new understanding on the local microbiome distribution and changes shall focus on oil or

hydrophobic fraction of the reservoir fluids (Kobayashi et al., 2012; Tang et al., 2012; Lenchi et al., 2013). Consequently, the alteration pattern of major methanogenesis in the aqueous and oil phases under different temperature conditions could be further delineated. A combination of methods, including synthesis and quantification of degradation intermediate (Bian et al., 2015), stable isotope labeling on tracing substrate transformation (Gray et al., 2011) and molecular analysis of the metabolically active microorganisms can advance the studies on anaerobic degradation and methanogenesis processes in reservoir systems.

3.3 Physicochemical influence and Taxa & Function profiles

Temperature is an important physical factor shaping the community structure of bacterial (anosim & adonis $P < 0.01$) and methanogenic communities (anosim $P < 0.05$ & adonis $P < 0.01$) of the samples in this study (Figure 4 and Table S7). Furthermore, a significant difference of taxa abundance among the three temperature categories for both bacterial and archaeal communities was evident by LEfSe analysis (Figure 5). For bacterial community, sample group (aqueous or oil phases from the same sample group), temperature (anosim & adonis $P < 0.01$), and pH (adonis $P < 0.05$) showed significant effects on separating samples into different categories. For archaeal community, significant differences among sample categories were detected with sample group and temperature (both adonis $P < 0.05$), pH and phase (anosim & adonis $P < 0.05$). For methanogenic archaeal community, significant differences among sample categories were detected with sample group and temperature (sample group: anosim & adonis $P < 0.05$; temperature: anosim $P < 0.05$ & adonis $P < 0.01$) (Table S7).

Nitrate concentration dissimilarity matrix was significantly associated with all unweighted/weighted UniFrac and Bray-Cutis matrices (all $P < 0.05$) for aqueous bacterial community (Table S8). Meanwhile, methanogenic archaeal community in oil phase was

significantly affected by reservoir depth, temperature, pH and water flooding operation years based on association analysis of weighted UniFrac matrix but not unweighted UniFrac matrix (all $P < 0.05$) (Table S8), indicating that it was the abundance difference of certain taxa affecting the compositional pattern. More detailed relationships between physicochemical factors and bacterial/archaeal/methanogenic communities are summarized in Tables S7 and S8.

Nitrate dissimilarity matrix was significantly correlated with dissimilarity matrices of all aqueous phase bacterial communities using both unweight/weight UniFrac and Bray-Curtis matrices (all $P < 0.05$) (Table S8). Nitrate is an important chemical used in the injection water to inhibit corrosion and maintain crude quality (Gao et al., 2013). Nitrate stimulates the growth of nitrate-reducing bacteria and inhibits the growth of SRB (Nemati et al., 2001; Gao et al., 2013). Consequently, nitrate injection shapes the microbial communities in petroleum reservoirs. Mn^{2+} and Mg^{2+} were shown to be strongly associated with bacterial community of aqueous phase based on Bray-Curtis matrix (both $P < 0.05$) (Table S8). Metal ions can be electron acceptor for direct or indirect hydrocarbon degradation under anaerobic conditions (Mbadinga et al., 2011). Metal reducers could also utilize electrons from syntrophic partners to further facilitate direct aromatic hydrocarbon degradation (Kunapuli et al., 2007).

Differences in taxa and function profiles in oil and aqueous phases were analyzed based on LEfSe and Tax4Fun (Figures S2 and S3, and Table S9). For bacterial communities, amino acid metabolism and xenobiotics biodegradation and metabolism were distributed more in oil phase, while carbohydrate metabolism was distributed more in aqueous phase. For archaeal communities, ubiquinone and other terpenoid quinone biosynthesis and butanoate and tryptophan metabolism together with other glycan degradation pathways were distributed more in oil phase, while folding, sorting and degradation as well as several infectious

diseases related pathways, such as *Salmonella* infection, tuberculosis, were distributed more in the aqueous phase. Since FTU values of archaeal communities (fraction of taxa that could be mapped to existing KEGG pathway) were unevenly distributed from sample to sample, the reliability of these functional predictions is in question (Table S9). Meanwhile, the database of Tax4Fun is far from completeness, due to the enormous amount of uncultured microorganism and their unknown genomes. To date, there are still very limited studies on investigating inhabiting preference of microbiome in petroleum reservoirs. Meanwhile, their functional contributions to hydrocarbon degradation and methanogenesis in both aqueous and oil phases remain elusive, which calls for further efforts on this point (Kryachko et al., 2012; Wang et al., 2014; Kobayashi et al., 2012).

4 Conclusions

A core bacterial microbiome containing a small proportion of OTUs but a relatively large proportion of sequences mediating hydrocarbon degradation and fermentation was revealed by analysis of oil reservoirs of different temperatures. The core and common bacterial microbiome of major biodegrading functions were shared among geographically and physicochemically different reservoirs. The different shifting patterns of the dominant methanogenesis pathway in aqueous and oil phase within samples of different temperature were featured. Factors of pH, temperature, phase conditions and nitrate, Mn^{2+}, and Mg^{2+} concentrations shaped the microbial compositional and functional profiles significantly. Moreover, biomarker groups of bacteria and archaea associated with different pH, temperature and phase conditions indicate major differences in function of amino acid metabolism, xenobiotics metabolism enriched in the oil phase, and carbohydrate metabolism enriched in the aqueous phase.

Author contributions. ZZ, BL, L-YW, B-ZM, HS, and J-DG conceived the project and designed experiments. ZZ, BL, and L-YW conducted the sampling, chemical and molecular experiments. J-FL, B-ZM, and J-DG managed sample collection, supervised on data interpretation. ZZ performed the original data analysis and drafted the original manuscript. All members contributed to refining the manuscript and approved the final version.

Competing interests. The authors declare no conflict of interest.

Acknowledgments. Dr. Xiangzhen Li's group at Chengdu Institute of Biology, Chinese Academy of Sciences was thanked for MiSeq sequencing efforts and related technical support. Ms. Kelly Lau was thanked for her supportive technician work. We thank the support from local administrative and technical staff in Shengli, Daqing, Huabei, Xinjiang Karamay, and Jiangsu Oilfields. This projected was supported by NSFC/RGC (grant no. N HKU718/11, J-DG) and NSFC (grant no. 41373070 and 41530318, B-ZM) and a postgraduate scholarship from HKU (ZZ).

References

Angel, R., Claus, P., and Conrad, R.: Methanogenic archaea are globally ubiquitous in aerated soils and become active under wet anoxic conditions, ISME J., 6, 847-862, 10.1038/ismej.2011.141, 2012.

Asshauer, K. P., Wemheuer, B., Daniel, R., and Meinicke, P.: Tax4Fun: predicting functional profiles from metagenomic 16S rRNA data, Bioinformatics, 31, 2882-2884, 10.1093/bioinformatics/btv287, 2015.

Bian, X.-Y., Mbadinga, S. M., Liu, Y.-F., Yang, S.-Z., Liu, J.-F., Ye, R.-Q., Gu, J.-D., and Mu, B.-Z.: Insights into the Anaerobic Biodegradation Pathway of *n*-Alkanes in Oil Reservoirs by Detection of Signature Metabolites, Sci. Rep., 5, 9801, 10.1038/srep09801, 2015.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Pena, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., Sevinsky, J. R., Tumbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T., Zaneveld, J., and Knight, R.: QIIME allows analysis of high-throughput community sequencing data, Nat. Methods, 7, 335-336, 10.1038/nmeth.f.303, 2010.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Huntley, J., Fierer, N., Owens, S. M., Betley, J., Fraser, L., Bauer, M., Gormley, N., Gilbert, J. A., Smith, G., and Knight, R.: Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms, ISME J., 6, 1621-1624, 10.1038/ismej.2012.8, 2012.

Dahle, H., Garshol, F., Madsen, M., and Birkeland, N.-K.: Microbial community structure analysis of produced water from a high-temperature North Sea oil-field, Antonie Van Leeuwenhoek, 93, 37-49, 10.1007/s10482-007-9177-z, 2008.
Daly, R. A., Borton, M. A., Wilkins, M. J., Hoyt, D. W., Kountz, D. J., Wolfe, R. A., Welch, S. A., Marcus, D. N., Trexler, R. V., and MacRae, J. D.: Microbial metabolisms in a 2.5-km-deep ecosystem created by hydraulic fracturing in shales, Nat. Microbiol., 1, 16146, 2016.
Edgar, R. C.: Search and clustering orders of magnitude faster than BLAST, Bioinformatics, 26, 2460-2461, 10.1093/bioinformatics/btq461, 2010.
Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R.: UCHIME improves sensitivity and speed of chimera detection, Bioinformatics, 27, 2194-2200, 10.1093/bioinformatics/btr381, 2011.
Gao, P., Li, G., Dai, X., Dai, L., Wang, H., Zhao, L., Chen, Y., and Ma, T.: Nutrients and oxygen alter reservoir biochemical characters and enhance oil recovery during biostimulation, World J. Microbiol. Biotechnol., 29, 2045-2054, 10.1007/s11274-013-1367-4, 2013.
Gao, P., Tian, H., Li, G., Sun, H., and Ma, T.: Microbial diversity and abundance in the Xinjiang Luliang long-term water-flooding petroleum reservoir, MicrobiologyOpen, 4, 332-342, 10.1002/mbo3.241, 2015a.
Gao, P. K., Li, G. Q., Tian, H. M., Wang, Y. S., Sun, H. W., and Ma, T.: Differences in microbial community composition between injection and production water samples of water flooding petroleum reservoirs, Biogeosciences, 12, 3403-3414, 10.5194/bg-12-3403-2015, 2015b.
Grabowski, A., Nercessian, O., Fayolle, F., Blanchet, D., and Jeanthon, C.: Microbial diversity in production waters of a low-temperature biodegraded oil reservoir, FEMS Microbiol. Ecol., 54, 427-443, 10.1016/j.femsec.2005.05.007, 2005.
Gray, N. D., Sherry, A., Grant, R. J., Rowan, A. K., Hubert, C. R. J., Callbeck, C. M., Aitken, C. M., Jones, D. M., Adams, J. J., Larter, S. R., and Head, I. M.: The quantitative significance of *Syntrophaceae* and syntrophic partnerships in methanogenic degradation of crude oil alkanes, Environ. Microbiol., 13, 2957-2975, 10.1111/j.1462-2920.2011.02570.x, 2011.
Guo, H., Liu, R., Yu, Z., Zhang, H., Yun, J., Li, Y., Liu, X., and Pan, J.: Pyrosequencing reveals the dominance of methylotrophic methanogenesis in a coal bed methane reservoir associated with Eastern Ordos Basin in China, Int. J. Coal Geol., 93, 56-61, 10.1016/j.coal.2012.01.014, 2012.
Head, I. M., Jones, D. M., and Larter, S. R.: Biological activity in the deep subsurface and the origin of heavy oil, Nature, 426, 344-352, 10.1038/nature02134, 2003.
Kobayashi, H., Endo, K., Sakata, S., Mayumi, D., Kawaguchi, H., Ikarashi, M., Miyagawa, Y., Maeda, H., and Sato, K.: Phylogenetic diversity of microbial communities associated with the crude-oil, large-insoluble-particle and formation-water components of the reservoir fluid from a non-flooded high-temperature petroleum reservoir, J. Biosci. Bioeng., 113, 204-210, 10.1016/j.jbiosc.2011.09.015, 2012.
Kryachko, Y., Dong, X., Sensen, C. W., and Voordouw, G.: Compositions of microbial communities associated with oil and water in a mesothermic oil field, Antonie Van Leeuwenhoek, 101, 493-506, 10.1007/s10482-011-9658-y, 2012.
Kunapuli, U., Lueders, T., and Meckenstock, R. U.: The use of stable isotope probing to identify key iron-reducing microorganisms involved in anaerobic benzene degradation, ISME J., 1, 643-653, 10.1038/ismej.2007.73, 2007.
Lee, S.-H., Park, J.-H., Kim, S.-H., Yu, B. J., Yoon, J.-J., and Park, H.-D.: Evidence of syntrophic acetate oxidation by *Spirochaetes* during anaerobic methane production, Bioresour. Technol., 190, 543-549, 10.1016/j.biortech.2015.02.066, 2015.
Lenchi, N., Inceoglu, O., Kebbouche-Gana, S., Gana, M. L., Lliros, M., Servais, P., and Garcia-Armisen, T.: Diversity of Microbial Communities in Production and Injection Waters of Algerian Oilfields Revealed by 16S rRNA Gene Amplicon 454 Pyrosequencing, PLoS One, 8, 10.1371/journal.pone.0066588, 2013.
Li, C.-Y., Zhang, D., Li, X.-X., Mbadinga, S. M., Yang, S.-Z., Liu, J.-F., Gu, J.-D., and Mu, B.-Z.: The biofilm property and its correlationship with high-molecular-weight polyacrylamide degradation in a water injection pipeline of Daqing oilfield, J. Hazard. Mater., 10.1016/j.jhazmat.2015.10.067, 2015.
Li, D., Midgley, D., Ross, J., Oytam, Y., Abell, G. J., Volk, H., Daud, W., and Hendry, P.: Microbial biodiversity in a Malaysian oil field and a systematic comparison with oil reservoirs worldwide, Arch. Microbiol., 194, 513-523, 10.1007/s00203-012-0788-z, 2012a.
Li, G., Gao, P., Wu, Y., Tian, H., Dai, X., Wang, Y., Cui, Q., Zhang, H., Pan, X., Dong, H., and Ma, T.: Microbial Abundance and Community Composition Influence Production Performance in a Low-Temperature Petroleum Reservoir, Environ. Sci. Technol., 48, 5336-5344, 10.1021/es500239w, 2014.

Li, H., Wang, X.-L., Mu, B.-Z., Gu, J.-D., Liu, Y.-D., Lin, K.-F., Lu, S.-G., Lu, Q., Li, B.-Z., Li, Y.-Y., and Du, X.-M.: Molecular detection, quantification and distribution of alkane-degrading bacteria in production water from low temperature oilfields, Int. Biodeterior. Biodegrad., 76, 49-57, 10.1016/j.ibiod.2012.06.007, 2013.
Li, W., Wang, L.-Y., Duan, R.-Y., Liu, J.-F., Gu, J.-D., and Mu, B.-Z.: Microbial community characteristics of petroleum reservoir production water amended with *n*-alkanes and incubated under nitrate-, sulfate-reducing and methanogenic conditions, Int. Biodeterior. Biodegrad., 69, 87-96, 10.1016/j.ibiod.2012.01.005, 2012b.
Liang, B., Wang, L.-Y., Mbadinga, S. M., Liu, J.-F., Yang, S.-Z., Gu, J.-D., and Mu, B.-Z.: *Anaerolineaceae* and *Methanosaeta* turned to be the dominant microorganisms in alkanes-dependent methanogenic culture after long-term of incubation, AMB Express, 5, 10.1186/s13568-015-0117-4, 2015.
Liu, Y., and Whitman, W. B.: Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea, Ann. N.Y. Acad. Sci., 1125, 171-189, 10.1196/annals.1419.019, 2008.
Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., Forster, W., Brettske, I., Gerber, S., Ginhart, A. W., Gross, O., Grumann, S., Hermann, S., Jost, R., Konig, A., Liss, T., Lussmann, R., May, M., Nonhoff, B., Reichel, B., Strehlow, R., Stamatakis, A., Stuckmann, N., Vilbig, A., Lenke, M., Ludwig, T., Bode, A., and Schleifer, K. H.: ARB: a software environment for sequence data, Nucleic Acids Res., 32, 1363-1371, 10.1093/nar/gkh293, 2004.
Magoc, T., and Salzberg, S. L.: FLASH: fast length adjustment of short reads to improve genome assemblies, Bioinformatics, 27, 2957-2963, 10.1093/bioinformatics/btr507, 2011.
Magot, M., Ollivier, B., and Patel, B. K. C.: Microbiology of petroleum reservoirs, Antonie Van Leeuwenhoek, 77, 103-116, 10.1023/a:1002434330514, 2000.
Mayumi, D., Mochimaru, H., Yoshioka, H., Sakata, S., Maeda, H., Miyagawa, Y., Ikarashi, M., Takeuchi, M., and Kamagata, Y.: Evidence for syntrophic acetate oxidation coupled to hydrogenotrophic methanogenesis in the high-temperature petroleum reservoir of Yabase oil field (Japan), Environ. Microbiol., 13, 1995-2006, 10.1111/j.1462-2920.2010.02338.x, 2011.
Mbadinga, S. M., Wang, L.-Y., Zhou, L., Liu, J.-F., Gu, J.-D., and Mu, B.-Z.: Microbial communities involved in anaerobic degradation of alkanes, Int. Biodeterior. Biodegrad., 65, 1-13, 10.1016/j.ibiod.2010.11.009, 2011.
Mbadinga, S. M., Li, K.-P., Zhou, L., Wang, L.-Y., Yang, S.-Z., Liu, J.-F., Gu, J.-D., and Mu, B.-Z.: Analysis of alkane-dependent methanogenic community derived from production water of a high-temperature petroleum reservoir, Appl. Microbiol. Biotechnol., 96, 531-542, 10.1007/s00253-011-3828-8, 2012.
Mesle, M., Dromart, G., and Oger, P.: Microbial methanogenesis in subsurface oil and coal, Res. Microbiol., 164, 959-972, 10.1016/j.resmic.2013.07.004, 2013.
Narihiro, T., and Sekiguchi, Y.: Oligonucleotide primers, probes and molecular methods for the environmental monitoring of methanogenic archaea, Microb. Biotechnol., 4, 585-602, 10.1111/j.1751-7915.2010.00239.x, 2011.
Nemati, M., Jenneman, G. E., and Voordouw, G.: Mechanistic study of microbial control of hydrogen sulfide production in oil reservoirs, Biotechnol. Bioeng., 74, 424-434, 10.1002/bit.1133, 2001.
Ollivier, B., and Alazard, D.: The oil reservoir ecosystem, in: Handbook of hydrocarbon and lipid microbiology, Springer-Verlag, Berlin; Heidelberg, Germany, 2259-2269, 2010.
Orphan, V. J., Taylor, L. T., Hafenbradl, D., and Delong, E. F.: Culture-dependent and culture-independent characterization of microbial assemblages associated with high-temperature petroleum reservoirs, Appl. Environ. Microbiol., 66, 700-711, 10.1128/aem.66.2.700-711.2000, 2000.
Orphan, V. J., Goffredi, S. K., Delong, E. F., and Boles, J. R.: Geochemical influence on diversity and microbial processes in high temperature oil reservoirs, Geomicrobiol. J., 20, 295-311, 10.1080/01490450303898, 2003.
Pham, V. D., Hnatow, L. L., Zhang, S., Fallon, R. D., Jackson, S. C., Tomb, J.-F., DeLong, E. F., and Keeler, S. J.: Characterizing microbial diversity in production water from an Alaskan mesothermic petroleum reservoir with two independent molecular methods, Environ. Microbiol., 11, 176-187, 10.1111/j.1462-2920.2008.01751.x, 2009.
Rideout, J. R., He, Y., Navas-Molina, J. A., Walters, W. A., Ursell, L. K., Gibbons, S. M., Chase, J., McDonald, D., Gonzalez, A., Robbins-Pianka, A., Clemente, J. C., Gilbert, J. A., Huse, S. M., Zhou, H.-W., Knight, R., and Caporaso, J. G.: Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences, PeerJ, 2, 10.7717/peerj.545, 2014.
Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., and Huttenhower, C.: Metagenomic biomarker discovery and explanation, Genome Biol., 12, R60, 10.1186/gb-2011-12-6-r60, 2011.

Sherry, A., Grant, R. J., Aitken, C. M., Jones, D. M., Head, I. M., and Gray, N. D.: Volatile hydrocarbons inhibit methanogenic crude oil degradation, Front. Microbio., 5, 10.3389/fmicb.2014.00131, 2014.
Takai, K., and Horikoshi, K.: Rapid detection and quantification of members of the archaeal community by quantitative PCR using fluorogenic probes, Appl. Environ. Microbiol., 66, 5066, 10.1128/aem.66.11.5066-5072.2000, 2000.
Tang, Y.-Q., Li, Y., Zhao, J.-Y., Chi, C.-Q., Huang, L.-X., Dong, H.-P., and Wu, X.-L.: Microbial Communities in Long-Term, Water-Flooded Petroleum Reservoirs with Different in situ Temperatures in the Huabei Oilfield, China, PLoS One, 7, 10.1371/journal.pone.0033535, 2012.
Voordouw, G., Armstrong, S. M., Reimer, M. F., Fouts, B., Telang, A. J., Shen, Y., and Gevertz, D.: Characterization of 16S rRNA genes from oil field microbial communities indicates the presence of a variety of sulfate-reducing, fermentative, and sulfide-oxidizing bacteria, Appl. Environ. Microbiol., 62, 1623-1629, 1996.
Wang, L.-Y., Gao, C.-X., Mbadinga, S. M., Zhou, L., Liu, J.-F., Gu, J.-D., and Mu, B.-Z.: Characterization of an alkane-degrading methanogenic enrichment culture from production water of an oil reservoir after 274 days of incubation, Int. Biodeterior. Biodegrad., 65, 444-450, 10.1016/j.ibiod.2010.12.010, 2011.
Wang, L.-Y., Duan, R.-Y., Liu, J.-F., Yang, S.-Z., Gu, J.-D., and Mu, B.-Z.: Molecular analysis of the microbial community structures in water-flooding petroleum reservoirs with different temperatures, Biogeosciences, 9, 4645-4659, 10.5194/bg-9-4645-2012, 2012.
Wang, L.-Y., Ke, W.-J., Sun, X.-B., Liu, J.-F., Gu, J.-D., and Mu, B.-Z.: Comparison of bacterial community in aqueous and oil phases of water-flooded petroleum reservoirs using pyrosequencing and clone library approaches, Appl. Microbiol. Biotechnol., 98, 4209-4221, 10.1007/s00253-013-5472-y, 2014.
Wang, Y., and Qian, P.-Y.: Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies, PLoS One, 4, 10.1371/journal.pone.0007401, 2009.
Warren, E., Bekins, B. A., Godsy, E. M., and Smith, V. K.: Inhibition of acetoclastic methanogenesis in crude oil- and creosote-contaminated groundwater, Biorem. J., 8, 1-11, 10.1080/10889860490465840, 2004.
Wright, A. D. G., and Pimm, C.: Improved strategy for presumptive identification of methanogens using 16S riboprinting, J. Microbiol. Methods, 55, 337-349, 10.1016/s0167-7012(03)00169-6, 2003.
Yamane, K., Maki, H., Nakayama, T., Nakajima, T., Nomura, N., Uchiyama, H., and Kitaoka, M.: Diversity and Similarity of Microbial Communities in Petroleum Crude Oils Produced in Asia, Biosci., Biotechnol., Biochem., 72, 2831-2839, 10.1271/bbb.80227, 2008.
Youssef, N., Elshahed, M. S., and McInerney, M. J.: Microbial Processes in Oil Fields: Culprits, Problems, and Opportunities, in: Adv. Appl. Microbiol., edited by: Laskin, A. L., Sariaslani, S., and Gadd, G., Advances in Applied Microbiology, 141-251, 2009.
Yu, Z., Garcia-Gonzalez, R., Schanbacher, F. L., and Morrison, M.: Evaluations of different hypervariable regions of archaeal 16S rRNA genes in profiling of methanogens denaturing by Archaea-specific PCR and gradient gel electrophoresis, Appl. Environ. Microbiol., 74, 889-893, 10.1128/aem.00684-07, 2008.
Zhou, Z., Chen, J., Cao, H., Han, P., and Gu, J.-D.: Analysis of methane-producing and metabolizing archaeal and bacterial communities in sediments of the northern South China Sea and coastal Mai Po Nature Reserve revealed by PCR amplification of *mcrA* and *pmoA* genes, Front. Microbio., 5, 789, 10.3389/fmicb.2014.00789, 2015.

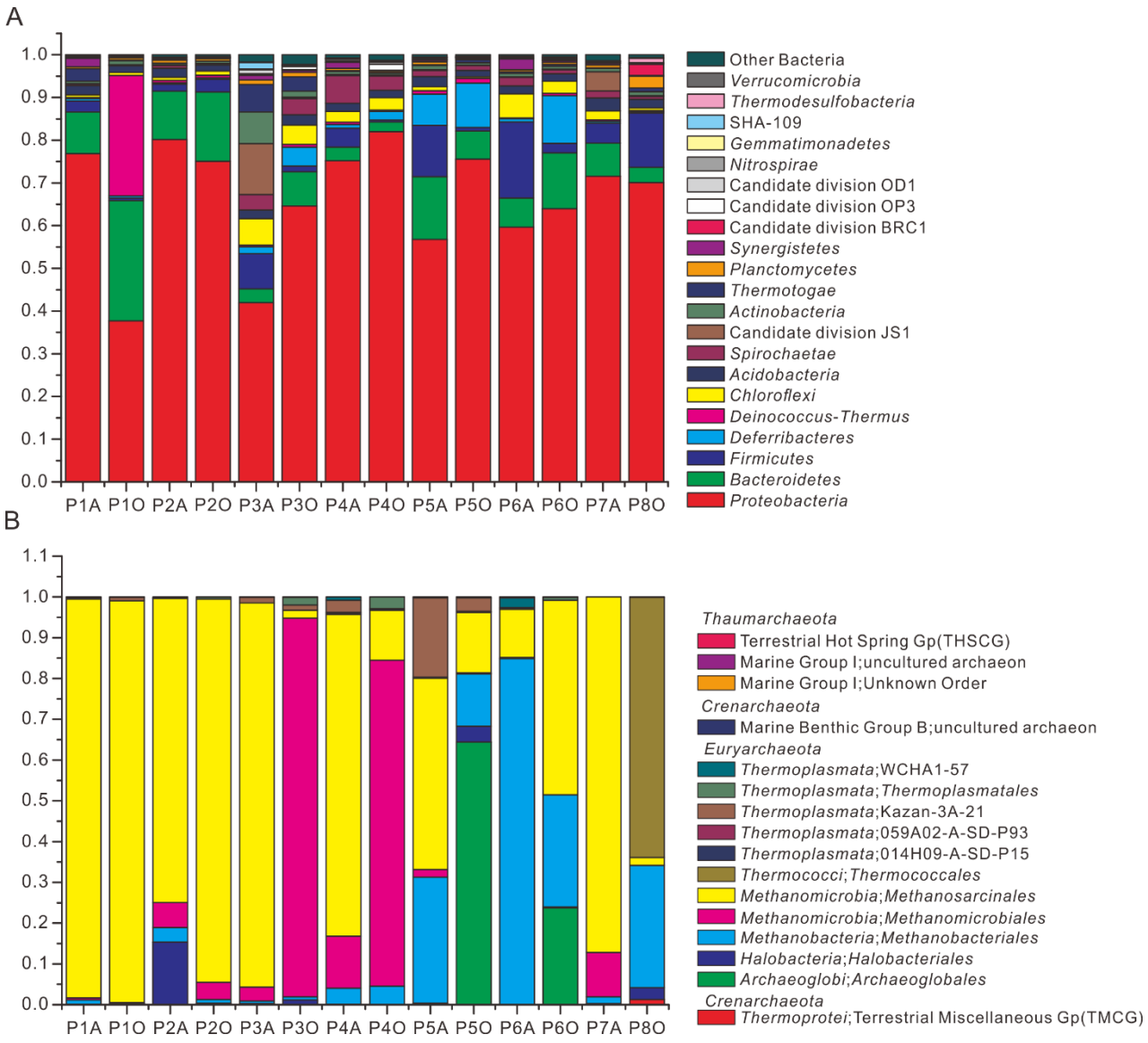

Figure 1. Relative abundance of bacteria (A) and archaea (B) from 14 aqueous and oil phase samples. Bacterial community was taxonomically assigned at the level of phylum. Those phyla with an average abundance of all samples below 0.1% were combined into "Other Bacteria" category. The archaeal community was taxonomically assigned at the level of class.

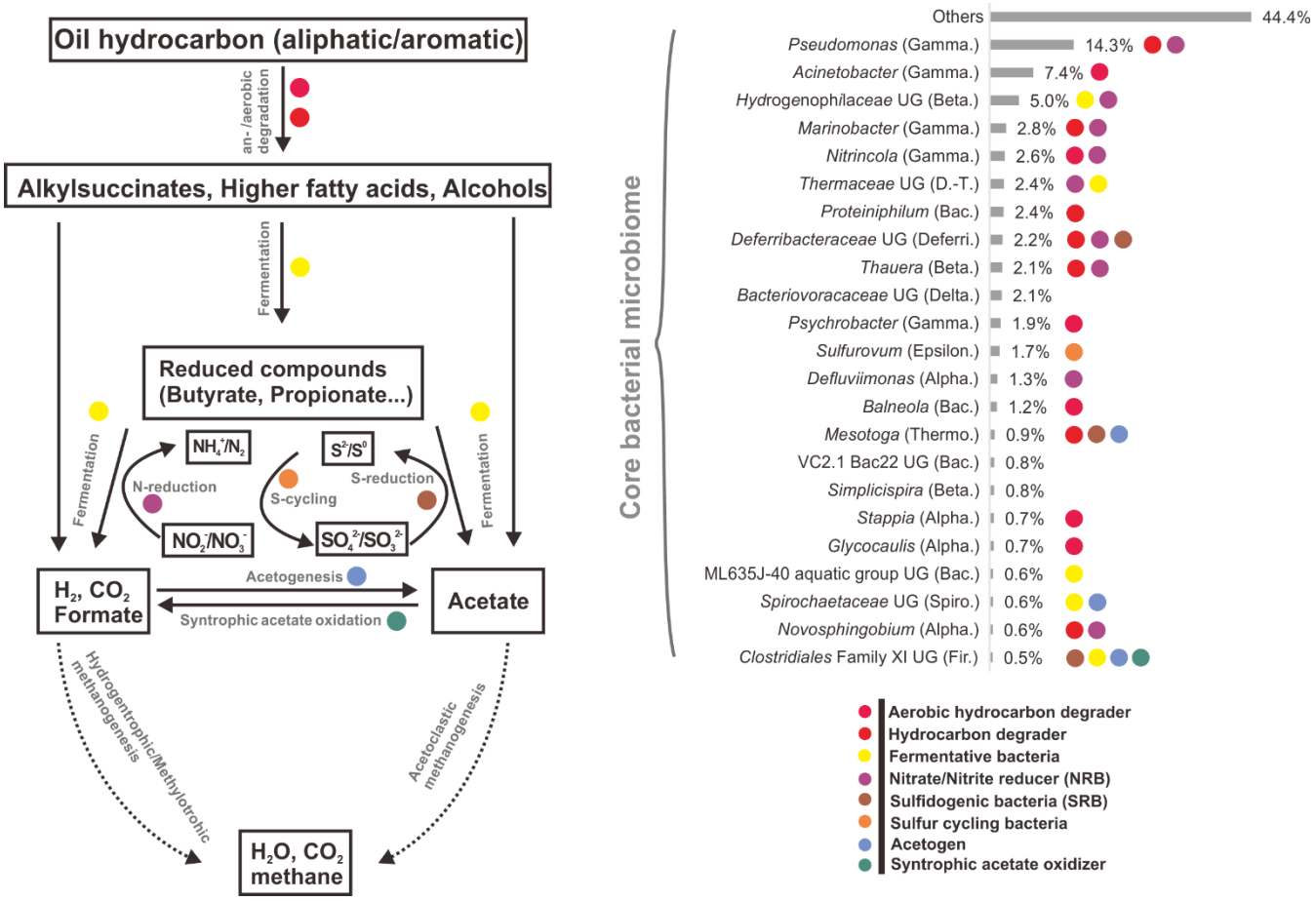

Figure 2. Average abundances and functional roles of core bacterial microbiome in all petroleum samples (including aqueous and oil phase samples). Core microbial OTUs (23 OTUs) were affiliated to the most abundant genera. The functional roles were assigned according to reference publications (Supplementary Material). Degradation pathway was modified from Mbadinga et al., 2011. Abbreviations: Alpha., Beta., Gamma., Delta., Epsilon (*Alpha-*, *Beta-*, *Gamma-*, *Delta-*, *Epsilonproteobacteria*), D.-T. (*Deinococcus-Thermus*), Bac. (*Bacteroidetes*), Deferri. (*Deferribacteres*), Delta. (*Deltaproteobacteria*), Thermo. (*Thermotogae*), Spiro. (*Spirochaetae*) and Fir. (*Firmicutes*). UG stands for uncultured genus.

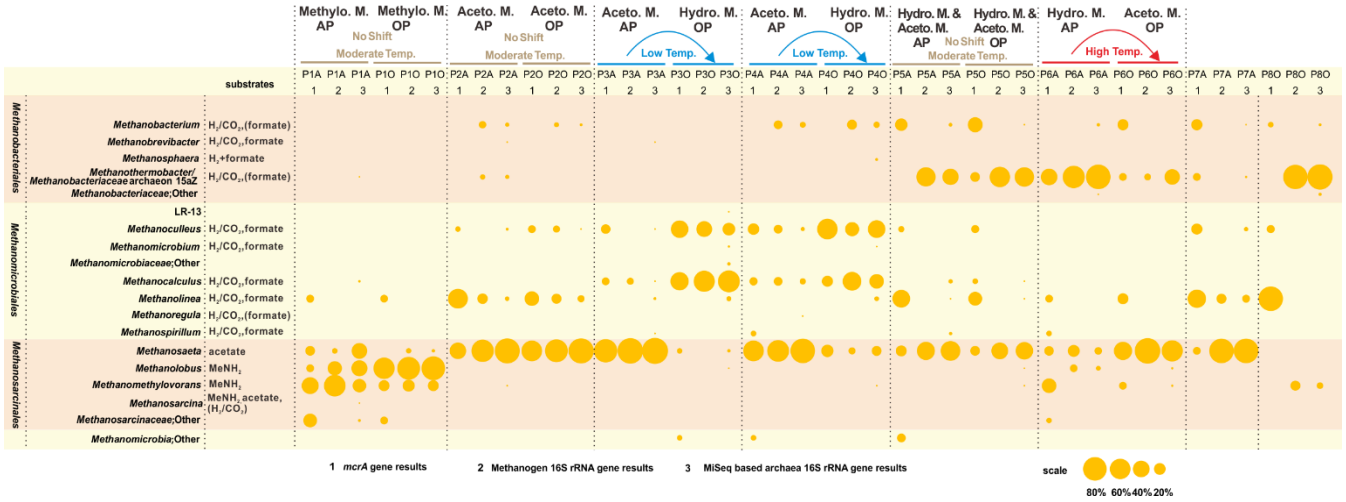

Figure 3. Bubble chart of the proportional composition of methanogens based on MiSeq sequencing results of archaeal 16S rRNA genes and clone sequencing results of *mcrA* gene and methanogenic 16S rRNA gene. The major substrate utilization properties were originated from Liu et al., 2008. $MeNH_2$ is methylamine and substrates in parentheses refer to being utilized by some but not all species. *Methanothermobacter thermautotrophicus* strain Delta H is the phylogenetically closest cultured clone to *Methanobacteriaceae* archaeon 15aZ. Thus, the combination of *Methanothermobacter* and *Methanobacteriaceae* archaeon 15aZ was used. Methanogenesis shift was based on the transition of major methanogenesis pathway. Abbreviations: Aceto. M., Hydro. M. and Methylo. M. (Acetoclastic methanogenesis, Hydrogenotrophic methanogenesis, and Methylotrophic methanogenesis).

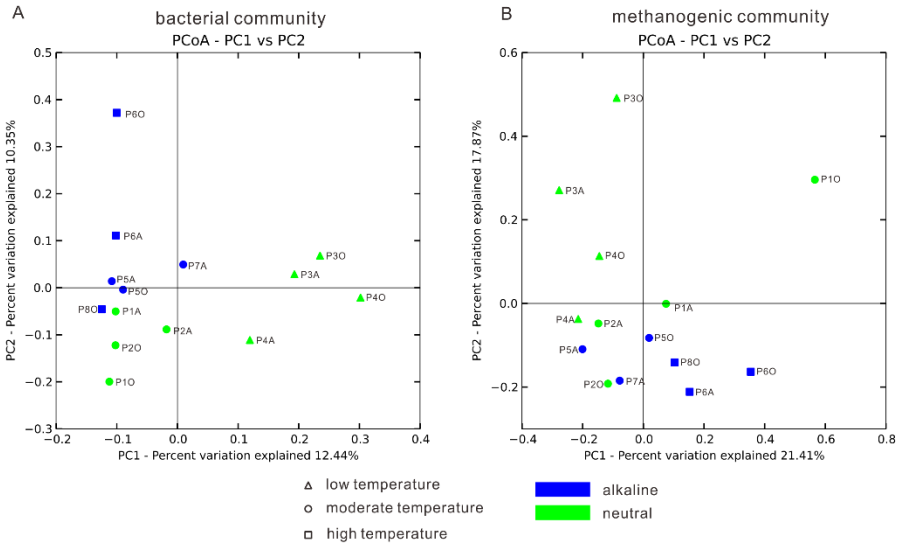

Figure 4. Principal Coordinate Analysis plot figures based on unweighted UniFrac matrices. Bacterial (A) and methanogenic (B) communities of 14 samples were separately analyzed to delineate the dissimilarity distances between each sample based on phylogenetic classification. The sample dots were categorized in terms of temperature (dot shape) and pH condition (color).

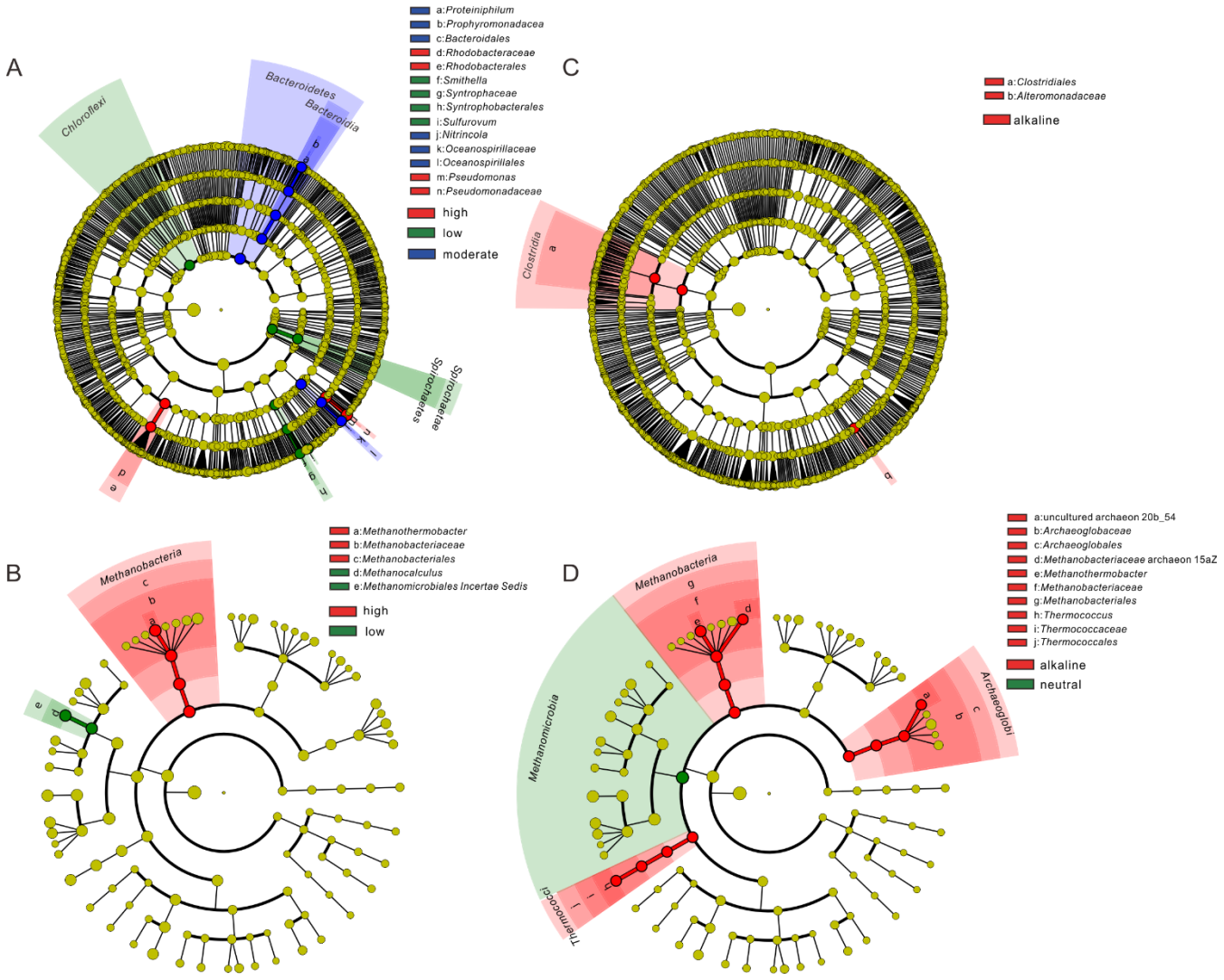

Figure 5. Cladogram based on LEfSe analysis results on bacteria (A, C) and archaea (B, D) in terms of temperature (A, B) and pH (C, D) categories. The taxonomic trees were generated from phylum to genus (inside to outside) in the hierarchical structure. Biomarker taxonomic levels were labeled in color which had at least logarithmic LDA score of 3.5. Pre-sample normalization was used to format the relative abundance data. All-against-all strategy was used in the multi-class analysis step.

Table 1. Reservoir geophysical properties and physicochemical properties of aqueous phase samples

Location name	Shengli-Zhan3-26	Huabei-B51	Xinjiang (Karamay)-Liuzhong	Xinjiang (Karamay)-Qixi	Jiangsu-Wei5	Jiangsu-Gao6	Daqing (Xingbei)-Xing4	Jiangsu-Qinying
Sample name	P1	P2	P3	P4	P5	P6	P7	P8
Depth (m)	~1300	~1101	~480	~1088	1018	1970	800-1200	2280-2524
Temp (°C)	63	45	21	32	53.5	82	49	88
pH	7.1	7.2	7.0	7.1	8.62	8.88	8.5	8.2
Effective porosity (%)	30	22.2	20.5	17.4	21.7	17.4	28.3	16.5
Average permeability ($\times 10^{-3}\mu m^2$)	800	12.6	466	274	91.8	94.5	481	22
Oil viscosity (mPa.s)	1720	402	417	44.8	292	8.4-83.9	13.86	98-188
Water flooding operation (years)	22	4	38	46	13	18	30	0
Cl^- (mg l^{-1})	3850	819	3864	2000	7563.01	1155.95	659.31	-
SO_4^{2-} (mg l^{-1})	2244	32.4	124.8	7.7	1921.06	156.81	333.29	-
PO_4^{3-} (mg l^{-1})	0.1	Nd	Nd	Nd	139.77	6.29	Nd	-
NO^{3-} (mg l^{-1})	Nd	Nd	34.1	1.4	Nd	0.44	0.45	-
Na^+ (mg l^{-1})	3313	1064	4196	5399	7363.52	1593.55	1363.43	-
K^+ (mg l^{-1})	94.2	22.3	35.1	45.6	9.36	3.26	3.55	-
Ca^{2+} (mg l^{-1})	195.6	53.0	103.3	128.2	17.42	7.24	41.94	-
Mg^{2+} (mg l^{-1})	46.1	17.6	44.7	64.0	10.27	2.25	11.02	-
Mn^{2+} (mg l^{-1})	0.3	0.1	0.3	0.4	Nd	Nd	Nd	-
Acetate (mg l^{-1})	32	57.9	344	6.97	44.07	285.66	Nd	-

"Nd" means not detected. No aqueous phase was obtained from Jiangsu-Qinying reservoir sample.

Table 2. PCR primers and PCR settings

	Primers	Sequences	Primer final concentration	Template DNA quantity	References
mcrA gene primer	ME3MF&ME3MF-e	ME3MF: ATGTCNGGTGGHGTMGGSTTYAC, ME3MF-e: ATGAGCGGTGGTGTCGGTTTCAC; concentration ratio = 250:1	1 μM	10-20 ng	(Narihiro and Sekiguchi, 2011)
	ME2r'	TCATBGCRTAGTTDGGRTAGT	1 μM	10-20 ng	
PCR setting	95°C2 min; 35 × (95°C 30 s, 59.5°C 30 s and 72°C 40 s); 72°C 5 min and 4°C 2 min.				
methanogenic 16S rRNA gene primer	MetA86f	GCTCAGTAACACGTGG	300 nM	10-20 ng	(Wright and Pimm, 2003; Narihiro and Sekiguchi, 2011; Yu et al., 2008)
	ARC915r	GTGCTCCCCCGCCAATTCCT	300 nM	10-20 ng	
PCR setting	95°C 5 min; 35 × (95°C for 30 s, 57°C for 30 s, and 72°C for 1 min); 72°C 10 min, 4°C 2 min.				
Prokaryotic universal primer	515F	GTGCCAGCMGCCGCGGTAA (12 nt barcode added)	1 μM	10 ng	(Caporaso et al., 2012; Wang and Qian, 2009)
	909R	GGACTACHVGGGTWTCTAAT	1 μM	10 ng	
PCR setting	94°C 3 min; 30 × (94°C 40 s, 56°C 60 s and 72°C 60 s); 72°C for 10 min, and 4°C for 2 min				
Archaeal universal primer	Arch347F	GYGCASCAGKCGMGAAW (12 nt barcode added)	1 μM	10 ng	(Takai and Horikoshi, 2000)
	Arch806R	GGACTACVSGGGTATCTAAT	1 μM	10 ng	
PCR setting	95°C for 5 min; 32 × (95°C 30 s, 50°C 30 s and 72°C 40 s); 72°C 10 min, and 4°C 2 min				