Differences in microbial community composition between injection and production water samples of water-flooding petroleum reservoirs

Author: P. K. Gao[1, 2], G. Q. Li[1, 2], H. M. Tian[1, 2], Y. S. Wang[1, 2], H. W. Sun[3], T. Ma[1, 2]

Affiliations:

[1] Key Laboratory of Molecular Microbiology and Technology, Ministry of Education, Tianjin 300071, P. R. China

[2] College of Life Sciences, Nankai University, Tianjin 300071, P. R. China

[3] College of Environmental Science and Engineering, Nankai University, Tianjin 300071, P. R. China

Corresponding author. T. Ma. Mailing address: College of Life Sciences, Nankai University, Tianjin 300071, P.R. China. Tel/Fax: 86-22-23498185.

E-mail: tingma@nankai.edu.cn

Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Abstract. Microbial communities in injected water are expected to have significant influence on those of reservoir strata in long-term water-flooding petroleum reservoirs. To investigate the similarities and differences in microbial communities in injected water and reservoir strata, high-throughput sequencing of microbial partial 16S rRNA of the water samples collected from the wellhead and downhole of injection wells, and from production wells in a homogeneous sandstone reservoir and a heterogeneous conglomerate reservoir were performed. The results indicate that a small number of microbial populations are shared between the water samples from the injection and production wells in the sandstone reservoir, whereas a large number of microbial populations are shared in the conglomerate reservoir. The bacterial and archaeal communities in the reservoir strata have high concentrations, which are similar with those in the injected water. However, microbial population abundance exhibited large differences between the water samples from the injection and production wells. The number of shared populations reflects the influence of microbial communities in injected water on those in reservoir strata to some extent, and show strong association with the unique variation of reservoir environments.

Keywords 16S rRNA • Microbial community• Pyrosequencing • Miseq • MEOR

1 **Introduction**

Water-flooding is an efficient and least expensive oil recovery process that is employed worldwide. Water-flooding is believed to be a continued reinoculation of reservoir with surface microorganisms. After long-term water-flooding, microbial populations possessing exceptional survival abilities in injected water are expected to flow into oil-bearing strata, in where, exogenous and indigenous microbial populations form a new complex ecosystem (Zhang et al., 2012a). When injecting nutrients and air into reservoir strata, these microbial populations can be stimulated, and produce metabolites, such as polysaccharide, surfactants, acid, alcohol, and biogas. Because these metabolites can improve reservoir properties by blocking preferred water flow paths, lowering interfacial tension between brines and the oil phase, and decreasing oil viscosity (Youssef et al., 2009), microbial enhanced oil recovery (MEOR) have been applied to petroleum industry, and is currently studied extensively (Abdel-Waly, 1999;Zhang et al., 2012b;Bao et al., 2009;Gao et al., 2013;Li et al., 2014).

Microbial populations inhabiting petroleum reservoirs play critical roles in the microbial enhancing of the oil recovery process. As a result, an increasing number of studies, especially those based on culture-independent methods, have been performed, and broaden our knowledge of microbial diversity in oil reservoirs (Al-Bahry et al., 2013;Kumaraswamy et al., 2011;Lenchi et al., 2013;Okoro et al., 2014;Wang et al., 2012). However, to our best knowledge, detailed studies about the effects of the microbial communities in injected water on those in reservoir strata remains poorly understood. Based on the 16S rRNA gene clone library method, several studies to date have suggested that despite being flooded by the same injected water, there is a

significant difference in the communities between each production well (Tang et al., 2012;Ren et al., 2011). Zhang et al. compared microbial communities in samples of injection and production wells from reservoirs with different in situ temperatures, and pointed that the effects of microorganisms in the injected waters on microbial community compositions in produced waters are strong associated with reservoir temperature (Zhang et al., 2012a). However, because of the low throughput of the clone library method, many infrequent microbial taxa may not be detected, making it difficult to compare microbial communities in detail.

If microbial populations in injected water can flow into reservoir strata and reach production wells, is the microbial community in the injected water expected to have a similar community composition to those in the production wells? If there is a large difference in community composition, what is the difference and how many microbial populations are shared? To explore these issues, we investigated the microbial populations and their abundance in injection and production wells in a homogeneous sandstone petroleum reservoir with a permeability of 522×10^{-3} μm^2 and inter-well spacing of 300–425 m using 16S rRNA pyrosequencing and real-time fluorescent quantitative PCR (qPCR). At the same time, we analyzed microbial communities in water samples collected from the wellhead and downhole of injection wells, and from production wells in a heterogeneous conglomerate water-flooding petroleum reservoir with a permeability of 362×10^{-3} μm^2 and inter-well spacing of 100–150 m. High-throughput sequencing provides the opportunity to compare microbial populations with unprecedented levels of coverage and detail. The variation in permeability, inter-well spacing, and heterogeneity of the reservoirs is benefit for exploring the influence of reservoir physical properties on microbial distribution in

injected water and reservoir strata. The similarity among microbial communities was investigated using hierarchical clustering and Principal Coordinate Analysis. Microbial populations were also clustered according to injection and production wells to highlight the populations that showed the highest variability.

2 Materials and methods

2.1 Sampling locations

The Lu and Liu field block reservoirs are located in the Xinjiang Oil Field, in the Junggar Basin of Xinjiang Uygur Autonomous Region, Northwest China. The Lu field block is a homogeneous sandstone reservoir that has been water-flooded since 2001. The depth of the sampling horizon is approximately 1200 m with a temperature of 37^{o}C. The porosity of the reservoir is 29.9%, with an average permeability of 522×10^{-3} μm^{2}. The density of the crude oil is 0.846 g/cm^{3}, with an oil viscosity of 18 mPa•s. In the investigated well group (an injection well and four production wells), injection well Lu3084, located in the center of the production wells, has a direct influence on the neighboring producers, with inter-well distances of 300–425 m. The Liu field block is a heterogeneous conglomerate reservoir that has been water-flooded for approximately 30 years. The depth of the block horizon is approximately 1088 m, with a temperature of 22.6^{o}C. The porosity of the reservoir is 18.96 %, with an average permeability of 362×10^{-3} μm^{2}. The oil density is 0.912 g/cm^{3}, with an oil viscosity of 80 mPa•s. The selected well group includes two injection and three production wells, with an inter-well spacing of 100–150 m. The production well T90 is located at the center of injection wells T86 and T93, while production wells T95 and T96 are located at the edge of the field block and are mainly flooded by injection well T93 (Fig. 1).

The concentrations of potential nutrient factors, including crude oil properties, total nitrogen (TN), total phosphorus (TP), and ion concentration of formation brines, are listed in Table 2. The differences in geochemical parameters between crude oil samples from the two blocks are not obvious, indicating similar oil formation characteristics and maturity. The crude oil in both blocks had a higher content of saturates and aromatics, which favor the growth of hydrocarbon-degrading bacteria (HDB), and some anaerobes, such as sulfate-reducing bacteria. The cations and anions of the water samples in the two blocks were similar, with lower nitrogen and phosphorus content, which are essential for the survival and growth of microorganisms. The lack of nitrogen and phosphorus implies a low metabolism level of microorganisms in reservoir.

2.2 Water samples collection and DNA extraction

Based on tracer techniques, the time interval for injected water to flow from an injection well into neighboring production wells was approximately 30–45 days in the sandstone reservoir, and 7–10 days in the conglomerate reservoir (data provided by the Xinjiang Oil Field Company). Injected water from the sandstone reservoir was collected on three occasions every 15 days between October 2012 and November 2012, and the produced water samples (from the reservoir strata) were collected along with the second injected water sample on three occasions at a 30-day interval. All the injected and produced water samples were collected randomly from sampling valves located on the wellhead. In the conglomerate reservoir, the injected water samples were collected in November 2011 from the wellhead and the zone close to downhole (obtained by backflow, that is, the injected water flowed upward through the injection

well) of the injection wells. Seven days later, the produced water samples were collected from neighboring production wells on three occasions at a 7-day interval. The collected water samples were completely filled into 15 L sterilized plastic bottles, which were immediately capped and sealed to avoid contamination and oxygen intrusion.

For DNA extraction, the residual oil was first removed by heating the sample to 60°C for 30 min and by phase separation in sterilized separatory funnels. Microbial cells were then collected from 5 L of each water sample by centrifugation at 4°C for 15 min at 10,000 × g in a high-speed centrifuge (Beckman, CA 92821, USA). The cell deposits collected from the same sampling location were mixed and resuspended with TE buffer (Tris 80 mM, EDTA 40 mM, pH 8.0), and then lysed using a mini bead-beater (BioSpec, Bartlesville, OK 74005, USA) at 200 rpm for 1 min at room temperature with 0.1 mm glass beads. After bead beating, lysozyme was added (final concentration of 1 mg/ml), and the samples were incubated at 37°C for 1 h. Following the lysozyme treatment, 120 μL sodium-dodecyl sulphate (20% SDS, W/V) was added and the samples were incubated at 65°C for 60 min. Total genomic DNA was then extracted from the suspension solution using an AxyPrepTM Genomic DNA miniprep kit (Axygen Biosciences, Tewksbury, MA 01876, USA) according to the manufacturer's instructions and stored at -80°C for subsequent study.

2.3 Pyrosequencing of partial 16S rRNA genes and sequence analysis

Broadly conserved primers 27F (5'-AGA GTT TGA TCC TGG CTC AG-3') and 533R (5'-TTA CCG CGG CTG CTG GCA C-3') were used to amplify the bacterial 16S rRNA gene, while primers 344F (5'-ACG GGG YGC AGC AGG CGC GA-3')

and 915R (5’-GTG CTC CCC CGC CAA TTC CT-3’) were used to amplify the archaeal 16S rRNA gene. PCR reactions were performed following the protocol described in the Supporting Information (SI). Replicate PCR products of the same sample were mixed in a PCR tube. Then, the amplicons from each reaction mixture were pooled in equimolar ratios based on concentration and subjected to emulsion PCR to generate amplicon libraries. Amplicon pyrosequencing was performed on a Roche Genome Sequencer GS FLX+ platform at Majorbio Bio-Pharm Technology, Shanghai, China.

Sequences generated from pyrosequencing were analyzed using default settings in the open source software package mothur (Schloss et al., 2009). The detailed process is described in the SI. Alpha diversity analyses, including rarefaction and computation of the Shannon, Simpson, Chao1 metric, and phylogenetic diversity (PD), were used to assess biodiversity. The similarity among microbial communities was determined using UniFrac analysis in which weighted and unweighted Principal Coordinate Analysis (PCoA) were performed based on OTUs abundance or phylogenetic relationships. Specific differences in community composition of samples were visualized using heatmaps, ggplot, and Venn diagrams using the R software package.

2.4 Miseq-sequencing of partial 16S rRNA genes and sequence analysis

The bacterial and archaeal 16S rRNA gene V4 region (300–350 bp) were amplified using primer set 515f (GTG CCA GCM GCC GCG GTAA) and 806r (GGA CTA CHV GGG TWT CTA AT) with the protocol described by Caporaso et al. (Caporaso et al., 2011;Caporaso et al., 2012). A composite sample for sequencing was created by combining equimolar ratios of amplicons from the individual samples, followed by

gel purification and ethanol precipitation to remove any remaining contaminants and PCR artifacts. Amplicon sequencing was conducted on an Illumina MiSeq platform at Novogene Co., Beijing, China.

Pairs of reads from the original DNA fragments were merged using FLASH (Magoc and Salzberg, 2011). Sequences were then analyzed using the Quantitative Insights Into Microbial Ecology (QIIME) software package (Caporaso et al., 2010), and the UPARSE pipeline (Edgar, 2013). The detailed process is described in the SI. The similarity among microbial communities was determined using UniFrac analysis in which weighted PCoA was performed based on OTUs composition and phylogenetic relationships. Specific differences in community composition of samples were visualized using heatmaps, ggplot, and Venn diagrams using the R package.

2.5 Quantification of community abundance

Evaluation of community abundance by real-time fluorescent qPCR was performed using the 16S rRNA gene as a molecular marker. Reactions were performed using the FastStart Universal SYBR Green Master PCR mix in a Bio-Rad iQ5 Sequence detection system. The primer set 8F (5′-AGA GTT TGA T(CT)(AC) TGG CTC-3′)/338R (5′-GCT GCC TCC CGT AGG AGT-3′) were used to quantify bacterial community, while 806F (5′-ATT AGA TAC CCS BGT AGT CC-33′)/958R (5′-YCC GGC GTT GAM TCC AAT T-3′) were used to quantify archaeal community (Gittel et al., 2009). Ten-fold serial dilutions of a known copy number of plasmid DNA containing the target gene were subjected to real-time PCR in triplicate to generate an external standard curve. The PCR efficiency and correlation coefficients for the standard curves were higher than 95%, and R^2 values were greater than 0.99

for the curves. The specificity of the PCR amplification was determined by the melting curve. Gene copy numbers in unknown samples were determined based on standard curves.

2.6 Sequence accession numbers

The raw reads were deposited in the National Center for Biotechnology Information (BioProject ID: PRJNA246768, http://www.ncbi.nlm.nih.gov/bioproject/246768).

3 Results

3.1 Microbial community composition in the sandstone reservoir

Up to 4016–5060 bacterial and 2688–2857 archaeal sequences were obtained by pyrosequencing in the sandstone reservoir. These sequences were assigned into 249–538 bacterial and 45–130 archaeal OTUs at a 3% cutoff (Table S1). The individual rarefaction, Shannon, and Phylogenetic diversity curves tended to approach the saturation plateau (Fig. S1). The results of qPCR indicated that the copy number of bacterial 16S rRNA in water samples of the injection well was 8.25×10^{6} copies ml^{-1}, while 1.5×10^{6} to 2.75×10^{6} copies ml^{-1} in the water samples of the production wells. Comparing with the bacteria, the number of archaea was about one percent of the bacterial number, with 3.75×10^{4} 16S rRNA copies ml^{-1} in injection well and 8.5×10^{3} to 5.75×10^{4} copies ml^{-1} in the production wells.

Phylogenetic analysis indicated that the injected water (Lu3084) was dominated by *Proteobacteria* (50.43%), *Cyanobacteria* (15.51%), and *Chloroflexi* (9.12%). Among the *Proteobacteria*, *Betaproteobacteria* (20.42%) and *Alphaproteobacteria* (19.63%) were numerically dominant, while a small quantity of *Deltaproteobacteria* (5.49%),

Gamaproteobacteria (4.44%), and *Epsilonproteobacteria* (0.32%) were detected (Fig. 2a I). The produced water from Lu3073 was dominated by *Proteobacteria* (65.35%) *Spirochaetes* (13.38%), and *Bacteroidetes* (12.38%). *Gammaproteobacteria* (23.96%), *Deltaproteobacteria* (22.16%), *Alphaproteobacteria* (13.47%), and *Spirochaetes* (13.38%) dominated at class level (Fig. 2a I). In the produced water from Lu3095, Lu1039, and Lu2180, *Proteobacteria* composed 78.58%–95.75% of the bacterial communities. *Alphaproteobacteria* (15.43%, 26.77%, 53.54%), *Betaproteobacteria* (23.48%, 50.57%, 12.94%), and *Epsilonproteobacteria* (2.79%, 4.38%, 25.54%) were dominant (Fig. 2a I).

To archaeal communities, more than 95% of the archaeal sequences were assigned to *Methanobacteria*, *Methanococci*, and *Methanomicrobia* (Fig. 2a II). In the injected water, 87% sequences were classed into *Methanomicrobia*, and the dominant genera were *Methanosaeta* (42.39%), *Methanomethylovorans* (25.57%), and *Methanolobus* (10.96%). Among them, *Methanosaeta* uses only acetate to produce CH_4. *Methanomicrobia* accounted for 84.03% in the produced water of Lu1039, and *Methanolobus* (83.46%) and *Methanococcus* (11.23%) were the dominant genera. The archaeal communities were much more conserved in the produced water at Lu2180, Lu3073, and Lu3095, with *Methanococcus* accounting for 95.34%, 90.79%, and 86.79%, respectively. The *Methanolobus* and *Methanococcus* are methylotrophic and hydrogenotrophic methanogens.

3.2 Microbial community composition in the conglomerate reservoir

Between 52719 to 129106 16S rRNA gene sequences were analyzed and assigned to 2623 to 3414 genus-level OTUs (Table S2). In combination with the relative

abundance, the number of bacterial and archaeal sequences was calculated, with the number of sequences per sample ranging in size from 51273 to 128980 and 85 to 1445, respectively (Fig. S2). Based on the results of qPCR, the copy number of bacterial bacterial 16S rRNA in the water samples ranged from 1.5×10^7 to 6.5×10^7 copies ml^{-1}, while archaeal 16S rRNA ranged from 4.5×10^5 to 8.5×10^5 copies ml^{-1}.

In contrast to the sandstone reservoir, *Proteobacteria*, *Bacteroidetes*, *Firmicutes*, *Spirochaetes*, and *Synergistetes* were simultaneously detected in the water samples from the injection and production wells, composing 85.7%–94.1% of all bacterial communities. Similar to the sandstone reservoir, more *Proteobacteria* were detected in water samples from production wells. At the class level, *Gammaproteobacteria*, *Epsilonproteobacteria*, *Alphaproteobacteria*, *Betaproteobacteria*, *Deltaproteobacteria*, *Bacteroidia*, *Bacilli*, and *Clostridia* composed 74.5%–83.7% of the bacterial communities in the water samples from the injection and production wells (Fig. 2b I). The archaea were mainly assigned to *Methanomicrobia*, *Methanococci*, *Methanobacteria*, *Thaumarchaeota*, *Parvarchaea*, and *Thermoplasmata* (Fig. 2b II). Among them, *Methanobacteria*, *Methanococci*, and *Methanomicrobia* were simultaneously detected, and composed 64.3%–94.6% of the archaeal communities in the water samples from the injection and production wells. Compared with the injected water collected from the wellhead of the injection wells (T86-0 and T93-0), more *Methanomicrobia* were detected in the downhole of injection wells (T86-8 and T93-7) and production well T90. At genus level, *Methanocorpusculum*, *Methanococcus*, and *Methanocalculus* were dominant, accounting for 60.3–88.5% of the archaeal communities in the injection wells and production well T90. The three taxa can use H_2 and formate as carbon sources to

produce CH_4.

3.3 Shared microbial populations between injected water and reservoir strata

The shared microbial OTUs and genera between communities in the water samples from the injection and production wells were investigated using Venn diagrams, histograms, and heatmap. Based on the Venn diagrams, 16.3%–32.81% of bacterial OTUs and 13.73%–51.61% of archaeal OTUs were shared between the water samples from the injection well and those of the production wells in the sandstone reservoir (Fig. 3a). These shared bacterial OTUs accounted for 4.6–24.71% of the total bacterial community in water samples from injection well, and 43.23–76.18% in the production wells (Fig. 4a I). Furthermore, only 13 bacterial and 3 archaeal OTUs were shared by both the water samples from the injection and production wells (Fig. 3a). Accordingly, the shared genera only accounted for 2.26% of the bacterial community in the injected water, but dominated each production well (12.02%–36.5%; Fig. 4a II). Similar to the bacteria, three archaeal genera belonging to *Methanobacterium*, *Methanococcus*, and *Methanolobus* were simultaneously detected, comprising 13.58% of the total archaea in the water samples from the injection well, and 90.4%–96.89% in the water samples from the production wells (Fig. 4a III).

In the conglomerate reservoir, most of the OTUs and genera were simultaneously detected in the water samples from the injection and production wells (Fig. 3b and 4b). Similar with the sandstone reservoir, these shared populations accounted for a minor proportion of the communities in the water samples collected from the wellhead of injection wells, but dominated the water samples obtained from the downhole of injection wells and each production well (Fig.3b).

3.4 Microbial population distribution in injected water and reservoir strata

Microbial populations were clustered according to injection and production wells to highlight the populations that showed the most variability (Fig. 5). In the sandstone reservoir, more *Sphingomonas* and *Azospirillum* were detected in the water samples from the injection well, while *Arcobacter*, *Marinobacterium*, *Pseudomonas*, *Hyphomonas*, *Novispirillum*, *Proteiniphilum*, *Spirochaeta*, and *Rhizobium* were highly abundant in the water samples from the production wells. In the conglomerate reservoir, higher amounts of *Paracoccus*, *Bacillus*, *Ochrobactrum*, *Parabacteroides*, *Sphaerochaeta*, *Thauera*, *Halomonas*, and *Alcanivorax* were detected in the water samples from the injection well, while *Arcobacter*, *Marinobacterium*, *Pseudomonas*, *Bacteroides*, *Oleibacter*, *Marinobacter*, and *Shewanella* were dominant in the downhole of the injection and production wells. Among them, *Marinobacterium*, *Paracoccus*, *Ochrobactrum*, *Sphingomonas*, *Alcanivorax*, and *Azospirillum* are aerobic bacteria, while *Pseudomonas*, *Rhizobium*, *Arcobacter*, *Halomonas*, *Spirochaeta*, *Bacillus*, *Thauera*, *Halomonas*, and *Bacteroides* are microaerophilic bacteria, facultative anaerobes, or anaerobes.

To further investigate the microbial distribution in injected water and reservoir strata, hierarchical clustering and Unifrac PCoA were performed based on microbial OTUs abundance and phylogenetic relationships. In the sandstone reservoir, hierarchical clustering showed that the community in the water samples from the injection well was distinct from those of the production wells (Fig. S3). Weighted PCoA distinguished the bacterial community of the injected water from those of the production wells, while communities of the production wells were placed at a

comparatively decentralized position (Fig. 6a bacteria). Similar to the bacterial communities, hierarchical clustering and PCoA distinguished the archaeal community of the injected water from those of the production wells, whereas production wells were placed at a close proximity (Fig. 6a archaea). In the conglomerate reservoir, communities of water samples collected from the wellhead of injection wells clustered into a group in the PCoA plot, indicating that communities remained unchanged before injected water flowed into the injection wells (Fig. 6b). Communities in the water samples collected from the downhole of injection wells and neighboring production well T90 clustered into one group, while production well T95 and T96 clustered into another (Fig. 6b). This shows that the microbial community reassembled during the process of the injected water flowing into the reservoir strata and each production well.

4 Discussion

MEOR technique is generally classified into exogenous MEOR and indigenous MEOR (Youssef et al., 2009). The former includes injection of exogenous microorganisms and injection of ex-situ produced products into reservoirs to enhance oil recovery (Zobell, 1947). This is an effective way to quickly improve oil recovery. However, because of the sieve effect of strata on microbial cells, the injected microorganisms are generally difficult to migrate into reservoir strata (Youssef et al., 2009;Brown, 2010). Diverse microbial populations inhabit petroleum reservoirs. Among them, hydrocarbon-degrading bacteria (HDB), nitrate-reducing bacteria (NRB), sulfate-reducing bacteria (SRB), and methanogens are the most important functional populations in reservoir ecosystems (Youssef et al., 2009). When injecting

nutrients into reservoir, these microbial populations can be stimulated and produce biosurfactants, organic acids, and gas, which improve oil recovery (Belyaev SS et al., 1998). Comparing with exogenous MEOR, indigenous microorganisms are more adapt to the environmental conditions present in reservoirs. Additionally, because nutrients can be easier to migrate into reservoir strata, indigenous MEOR has higher oil displacement efficiency. Despite the validity in field trial, indigenous MEOR also has some limitations, in particular, the uneven oil displacement efficiency in different production wells in the same reservoir block. The community composition and diversification have been found to have a significant influence on oil displacement efficiency (Li et al., 2014). Therefore, it is needed to investigate the relationship between microbial communities in injected water and reservoir strata, because microbial communities in injected water are expected to flow into oil-bearing strata, and produce a significant influence on those of reservoir strata in long-term water-flooding petroleum reservoirs (Youssef et al., 2009;Dahle et al., 2008).

Molecular methods have been widely used to assess the microbial diversity of petroleum reservoirs. Compared to the traditional 16S rRNA gene clone library and sequencing, high-throughput sequencing has generated hundreds of thousands of short sequences, and significantly improved our ability to compare microbial populations with unprecedented levels of coverage and detail (Caporaso et al., 2012). In the conglomerate reservoir, Miseq-sequencing produced approximately 52719 to 129106 16S rRNA gene sequences. The sequencing reads was approximately 10–20 folds of those (obtained by pyrosequencing) of the sandstone reservoir, and 50–400 folds of

the 16S rRNA gene clone library (assuming 300 clones per library). We simultaneously sequenced the bacterial and archaeal V4 region of 16S rRNA gene, obtaining a total of 51273–128980 bacterial sequences per sample, but only 85–1445 archaeal sequences. This is consistent with the count result for archaea, which are about one percent of bacterial number. The result implies that deeper sequencing is needed for detecting rare archaeal populations using Miseq-sequencing based on bacterial and archaeal universal primer set 515f and 806r. In contrast, the bacterial and archaeal communities were sequenced independently using pyrosequencing in the sandstone and we obtained 4016–5060 bacterial and 2688–2857 archaeal sequences. The rarefaction curves, Shannon diversity estimates, and observed species, suggesting that this sequencing depth was enough for the investigation of the bacterial and archaeal communities.

If the microbial populations in the injected water could flow into the reservoir strata and reach the production wells along with the injected water, is the microbial community in the injected water expected to have a similar community composition with those in the production wells? In the homogeneous sandstone reservoir, we found the number of shared bacterial and archaeal populations between the water samples from the injection and production wells was different. As shown in Fig. 3a, 16.3%–32.81% of bacterial OTUs and 13.73%–51.61% of archaeal OTUs were shared between the water sample of injection well and those of production wells. It is reasonable to speculate that microbial populations in the injected water produce different levels of impact on those in production wells. Based on the previous research, the main reason may be the sieve effect that can be enhanced by the long inter-well spacing (Ren et al., 2011). Because of this effect on microbial cells

when injected fluid passes through a subsurface formation, it is more difficult for microbial cells to migrate in reservoir strata. In contrast, almost all OTUs and genera detected in the injected water were also observed in downhole of the injection and neighboring production wells in the heterogeneous reservoir. Compared with the sandstone reservoir, this reservoir has a similar permeability, but shorter inter-well spacing. It appears that most microbial populations in the injected water flowed into the oil-bearing strata and reached the production wells during water-flooding process. Additionally, despite lacking for sufficient nutrients, bacterial and archaeal communities in the reservoir strata have high concentrations, which are similar with those in the injected water. This phenomenon implies some correlations of microbial communities in injected water and reservoir strata. However, we appreciate that it is less rigorous to delineate the transport of microbial populations in the reservoir strata simply by detecting the shared microbial populations in the injection and production wells using 16S rRNA sequencing, because this method is not able to demonstrate whether the species detected in the produced water are the same ones as in the injected water. Therefore, labelled strains, such as ones containing green fluorescent protein, may be a suitable way to investigate microbial migration in petroleum reservoirs.

The number of shared microbial populations reflects the influence of microbial communities in injected water on those in reservoir strata to some extent, and show strong association with the unique variation of reservoir environments. Differ with the sandstone reservoir, a large number of microbial populations were shared between the water samples of injection wells and those of production wells in the conglomerate reservoir. However, the community structure, in particular, the abundance of the

shared OTUs and genera, exhibited a large difference between the injected water and reservoir strata. The environmental variables, such as salinity, pH, and nutrients, have been supposed to be the primary drivers for the community diversification (Kuang et al., 2013). However, few differences in cations and anions among the injected and produced water samples were observed. Petroleum reservoir represents an anaerobic environment with multiphase fluids of oil, gas and water. Therefore, except for the sieve effect of reservoir strata on microbial migration, dissolved oxygen, which is known to be strongly related to reservoir microbial growth and metabolism (Gao et al., 2013), may be the main factor influencing the community structures. In both the reservoirs, aerobic bacteria, including *Sphingomonas*, *Azospirillum*, *Paracoccus*, *Ochrobactrum*, *Alcanivorax*, and *Hydrogenophilaceae* were more frequently detected in water samples from the injection wells, while microaerophilic bacteria, facultative anaerobes, and anaerobes, including *Pseudomonas*, *Rhizobium*, *Arcobacter*, *Halomonas*, *Spirochaeta*, and *Bacteroides*, were found to have higher relative abundance in reservoir strata (Fig. 5).

Apart from the dissolved oxygen, crude oil, in particular, the saturated and aromatic hydrocarbon, may also strongly influence the microbial distribution in injected water and reservoir strata. Petroleum reservoirs represent oligotrophic environments. Although diverse microbial populations inhabit the reservoirs, only hydrocarbon-degrading bacteria and some anaerobes, such as sulfate-reducing bacteria, could grow with crude oil as carbon source. This is consistent with the observed results that more hydrocarbon-degrading bacteria, including *Marinobacterium*, *Pseudomonas*, *Rhizobium*, *Halomonas*, and *Oleibacter*, were detected in downhole of injection and production wells.

This study compared the differences in microbial community composition between injected water and reservoir strata using microbial genomes obtained from the aqueous phase. In fact, each component of the reservoir multiphasic fluid, including crude oil, gases, and insoluble particles, may act as an important habitat for microbial growth in addition to the water phase within the petroleum reservoir (Kryachko et al., 2012;Kobayashi et al., 2012). Recent research has also compared microbial communities in aqueous and oil phases of water-flooded petroleum reservoirs, and found that the oil phase also harbored a large number of microorganisms, with large differences in the bacterial community between the aqueous and oil phases of the reservoir fluid (Wang et al., 2014). Therefore, simultaneous analysis of DNA extracted from both aqueous and oil phases may provide a better understanding of the microbial communities in injection and production water samples.

In summary, this study investigated the relationship shared by microbial communities in injected water and reservoir strata in two long-term water-flooding reservoirs. However, the results cannot provide any reliable information on the indigenous microbial community. The indigenous microbial populations may be those in newly drilled wells without water-flooding in the same oil-bearing block. However, the two reservoirs have been water-flooded for decades. Due to the introduction of exogenous microorganisms in injected water and other sources of contaminations by enhanced oil recovery processes, determining whether a microorganism is indigenous to a petroleum reservoir become increasingly difficult. This study implies that the number of shared populations reflects the influence of microbial communities in injected water on those in reservoir strata to some extent, and show strong association with the

unique variation of reservoir environments. However, it cannot make a conclusion on the transport of microbial populations in the reservoir strata by detecting the shared microbial populations in injected water and produced water samples using 16S rRNA sequencing. To further investigate the relationship shared by microbial communities in injection and production wells, injecting labelled strains containing marked gene (e.g., green fluorescent protein coded gene) into reservoirs may bring novel insight and greater predictive power. Therefore, further research on microbial diversification and transferability as injected water flows into reservoir is needed. Solving these problems is significant to guide the application of MEOR approaches based on injecting nutrients or microbial populations into reservoirs.

5 Conclusions

Using high-throughput sequencing, this study revealed the similarities and differences in microbial communities in water samples from injection wells and those from reservoir strata in a homogeneous sandstone reservoir and a heterogeneous conglomerate reservoir. The bacterial and archaeal communities in the reservoir strata have high concentrations, which are similar with those in water samples from injection wells. However, microbial community compositions exhibited large differences between the water samples from injection and production wells. The number of shared populations reflects the influence of microbial communities in injected water on those in reservoir strata to some extent, and show strong association with the unique variation of reservoir environments. Additionally, aerobic bacterial populations were more frequently detected in water samples from injection wells, while microaerophilic bacteria, facultative anaerobes, and anaerobes dominated the reservoir strata.

Acknowledgments

This study was supported by the National High Technology Research and Development Program of China (Grant No. 2013AA064402), the National Natural Science Foundation of China (Grant No. 41373074), the Chinese Postdoctoral Science Foundation (Grant No. 2014M561175), and the Key Technology Research and Development Program of Tianjin (Grant No. 14ZCZDSY00016).

References

Abdel-Waly, A. A.: Laboratory study on activating indigenous microorganisms to enhance oil recovery, J. Can. Petrol. Technol., 38, 55-61, 1999.

Al-Bahry, S. N., Elshafie, A. E., Al-Wahaibi, Y. M., Al-Bemani, A. S., Joshi, S. J., Al-Maaini, R. A., Al-Alawi, W. J., Sugai, Y., and Al-Mandhari, M.: Microbial consortia in oman oil fields: a possible use in enhanced oil recovery, J. Microbiol. Biotechnol., 23, 106-117, 2013.

Bao, M. T., Kong, X. P., Jiang, G. C., Wang, X. L., and Li, X. M.: Laboratory study on activating indigenous microorganisms to enhance oil recovery in Shengli Oilfield, J Pet Sci Eng, 66, 42-46, 2009.

Belyaev SS, Borzenkov IA, and IF, G.: Activation of the geochemical activity of strata microflora as basis of a biotechnology for enhancement, Microbiology (Russ. Acad. Sci.), 67, 708~714, 1998.

Brown, L. R.: Microbial enhanced oil recovery (MEOR), Curr. Opin. Biotechnol., 13, 316–320, 2010.

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., Turnbaugh, 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, 2010.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., Fierer, N., and Knight, R.: Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample, Proc Natl Acad Sci U S A, 108 Suppl 1, 4516-4522, 2011.

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, 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.

Edgar, R. C.: UPARSE: highly accurate OTU sequences from microbial amplicon reads, Nat Meth, 10, 996-998, 2013.

Gao, P. K., Li, G. Q., Dai, X. C., Dai, L. B., Wang, H. B., Zhao, L. X., Chen, Y. H., and Ma, T.: Nutrients and oxygen alter reservoir biochemical characters and enhance oil recovery during biostimulation, World. J. Microb. Biot., 29, 2045–2054, 2013.

Gittel, A., Sørensen, K. B., Skovhus, T. L., Ingvorsen, K., and Schramm, A.: Prokaryotic Community Structure and Sulfate Reducer Activity in Water from High-Temperature Oil Reservoirs with and without Nitrate Treatment, Appl. Environ. Microbiol., 75, 7086–7096, 2009.

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, 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, 2012.

Kuang, J.-L., Huang, L.-N., Chen, L.-X., Hua, Z.-S., Li, S.-J., Hu, M., Li, J.-T., and Shu, W.-S.: Contemporary environmental variation determines microbial diversity patterns in acid mine drainage, Isme J, 7, 1038-1050, 2013.

564 Kumaraswamy, R., Ebert, S., Gray, M. R., Fedorak, P. M., and Foght, J. M.: Molecular- and
565 cultivation-based analyses of microbial communities in oil field water and in microcosms amended
566 with nitrate to control H2S production, Appl. Microbiol. Biotechnol., 89, 2027–2038, 2011.
567 Lenchi, N., İnceoğlu, Ö., Kebbouche-Gana, S., Gana, M. L., Llirós, M., Servais, P., and
568 García-Armisen, T.: Diversity of Microbial Communities in Production and Injection Waters of
569 Algerian Oilfields Revealed by 16S rRNA Gene Amplicon 454 Pyrosequencing, Plos One, 8, e66588,
570 2013.
571 Li, G. Q., Gao, P. K., Wu, Y. Q., Tian, H. M., Dai, X. C., Wang, Y. S., Cui, Q. S., Zhang, H. Z., Pan, X.
572 X., Dong, H. P., and Ma, T.: Microbial Abundance and Community Composition Influence Production
573 Performance in a Low-Temperature Petroleum Reservoir, Environ Sci Technol, 48, 5336-5344, 2014.
574 Magoc, T., and Salzberg, S. L.: FLASH: fast length adjustment of short reads to improve genome
575 assemblies, Bioinformatics, 27, 2957-2963, 2011.
576 Okoro, C., Smith, S., Chiejina, L., Lumactud, R., An, D., Park, H. S., Voordouw, J., Lomans, B. P., and
577 Voordouw, G.: Comparison of microbial communities involved in souring and corrosion in offshore and
578 onshore oil production facilities in Nigeria, J Ind Microbiol Biotechnol, 41, 665-678, 2014.
579 Ren, H. Y., Zhang, X. J., Song, Z. Y., Rupert, W., Gao, G. J., Guo, S. X., and Zhao, L. P.: Comparison
580 of microbial community compositions of injection and production well samples in a long-term
581 water-flooded petroleum reservoir, Plos One, 68, e23258, 2011.
582 Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A.,
583 Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J.,
584 and Weber, C. F.: Introducing mothur: open-source, platform-independent, community-supported
585 software for describing and comparing microbial communities, Appl. Environ. Microbiol., 75,
586 7537–7541, 2009.
587 Tang, Y. Q., Li, Y., Zhao, J. Y., Chi, C. Q., Huang, L. X., Dong, H. P., and Wu, X. L.: Microbial
588 communities in long-term, water-flooded petroleum reservoirs with different in situ temperatures in the
589 Huabei Oilfield, China, Plos One, 7, e33535, 2012.
590 Wang, L. Y., Duan, R. Y., Liu, J. F., Yang, S. Z., Gu, J. D., and Mu, B. Z.: Molecular analysis of the
591 microbial community structures in water-flooding petroleum reservoirs with different temperatures,
592 Biogeosciences, 9, 5177–5203, 2012.
593 Wang, L. Y., Ke, W. J., Sun, X. B., Liu, J. F., Gu, J. D., and Mu, B. Z.: Comparison of bacterial
594 community in aqueous and oil phases of water-flooded petroleum reservoirs using pyrosequencing and
595 clone library approaches, Appl Microbiol Biot, 98, 4209-4221, 2014.
596 Youssef, N., Elshahed, M. S., and McInerney, M. J.: Microbial processes in oil fields: culprits,
597 problems, and opportunities, Adv. Appl. Microbiol., 66, 141–251, 2009.
598 Zhang, F., She, Y. H., Chai, L. J., Banat, I. M., Zhang, X. T., Shu, F. C., Wang, Z. L., Yu, L. J., and Hou,
599 D. J.: Microbial diversity in long-term water-flooded oil reservoirs with different in situ temperatures in
600 China, Sci Rep, 2, 760, 10.1038/srep00760, 2012a.
601 Zhang, F., She, Y. H., Li, H. M., Zhang, X. T., Shu, F. C., Wang, Z. L., Yu, L. J., and Hou, D. J.: Impact
602 of an indigenous microbial enhanced oil recovery field trial on microbial community structure in a high
603 pour-point oil reservoir, Appl. Microbiol. Biotechnol., 95, 811–821, 2012b.
604 Zobell, C.: Bacterial release of oil from sedimentary materials, Oil Gas Journal, 46, 62-65, 1947.
605
606

Figure captions

Fig. 1. Schematic diagram showing (a) the distribution of the injection and production wells of the Liu and Lu field blocks, and (b) the wellhead and downhole of injection and production wells, and the location where the water samples were collected. The injected water samples were collected from the wellhead and the zone close to downhole (obtained by backflow, that is, the injected water flowed upward through the injection well) of the injection wells. The water samples in reservoir strata were collected from the wellhead of production wells. T86, T93 and Lu3084 are the selected injection wells, and T90, T95, T96, Lu1039, Lu2180, Lu3073 and Lu3095 are the production wells.

Fig. 2. Relative proportion of microbial taxa at class level in the injected and produced water samples. (a) Sandstone and (b) conglomerate reservoirs. I: bacterial taxa at class level; II: archaeal taxa at class level; Lu3084, T86-0, and T93-0: water samples from the well head of the injection wells; T86-8 and T93-7: water samples from downhole of the injection wells; T90, T95, T96, Lu1039, Lu2180, Lu3073, and Lu3095: water samples from the well head of the production wells.

Fig. 3. Venn diagrams of the bacterial and archaeal OTUs in the injection and production wells. (a) Sandstone and (b) conglomerate reservoirs. I: bacterial OTUs; II: archaeal OTUs. Venn diagrams indicate the shared microbial OTUs between communities in the injected and produced water samples.

Fig. 4. Comparison of shared microbial genera between the injection and production wells. a-I: pairwise comparison between injection and production wells in the sandstone reservoir; a-II: shared bacterial genera in the injection and production wells; a-III: shared archaeal genera in the injection and production wells; b-I and

b-II: comparison between injection and production wells on the conglomerate reservoir; and b-III: dominant shared bacterial genera in the conglomerate reservoir.

Fig. 5. Genera showing the most variability in the injected water and production wells. (a) Sandstone and (b) conglomerate reservoirs. The black-bordered box indicates the genera most detected in the production wells.

Fig. 6. Principal coordinate analysis of microbial communities used to investigate the microbial distribution in injected water and reservoir strata. (a) Sandstone and (b) conglomerate reservoirs. Bacteria: bacterial community distribution; Archaea: archaeal community distribution. Sample points that are close together are more similar in community composition than those that are far apart. The arrows in panel (b) indicate the community succession during the process of the injected water flowing into the injection wells and the neighbouring production wells.

644 **Table 1** Reservoir characteristics of Lu and Liu field block

Reservoir characteristics	Lu field block	Liu field block
Oil Reservoir		
Formation lithology	Sandstone	Conglomerate
Average depth (m)	1200	1088
Pressure (MPa)	10.2	7.2
Stratal temperature (^{o}C)	37	22.6
Average water content, %	80.8%	86.8%
Interwell distances, m	300-425	100-150
Average permeability, μm^2	522×10^{-3}	362×10^{-3}
Effective porosity, %	29.9	18.96
Water flooding (yr)	13	30
Crude oil properties		
Density (g/cm^3)	0.846	0.912
Viscosity in situ (mPa•s)	18	80.0
Saturates (%)	71.29	61.94
Aromatic (%)	14.85	11.24
Resin (%)	5.94	18.85
Asphalte (%)	5.94	7.97

645

646 **Table 2** Chemical properties of the water samples obtained from Lu and Liu field
647 block (Unit: mg/L)

Samples	Lu field block					Liu field block						
	3084	1039	2180	3073	3095	T86–0	T86–8	T93–0	T93–7	T90	T95	T96
Salinity	10850	11690	11170	10545	11102	10101	11313	11399	13991	13203	8997	9710
Nitrogen	15.1	11.5	10.6	12.7	11.6	7.6	6.8	10.2	11.5	5.7	6.3	8.5
Phosphorus	20.2	19.1	17.5	18.8	19.5	16.8	15.1	22.6	21.2	18.5	16.2	12.1
Na^+ K^+	4525	4803	4565	4309	4487	3364	3630	3802	4349	4014	3097	3139
Mg^{2+}	21.7	32.1	31.6	26.0	28.8	33.1	63.1	28.4	63.1	68.1	17.5	50.2
Ca^{2+}	191.3	281.9	284.7	181.6	216.4	70.2	77.9	72.8	78.0	96.2	86.4	108.7
Cl^-	5640	6125	5820	5160	5850	3010	3630	2922	3453	3099	3816	3406
SO_4^{2-}	116.2	14.04	4.86	23.13	8.86	483.4	89.4	523	362.4	89.9	65.6	165.1
HCO_3^-	356	434	464	846	511	3140	3823	4052	5687	5837	1915	2841

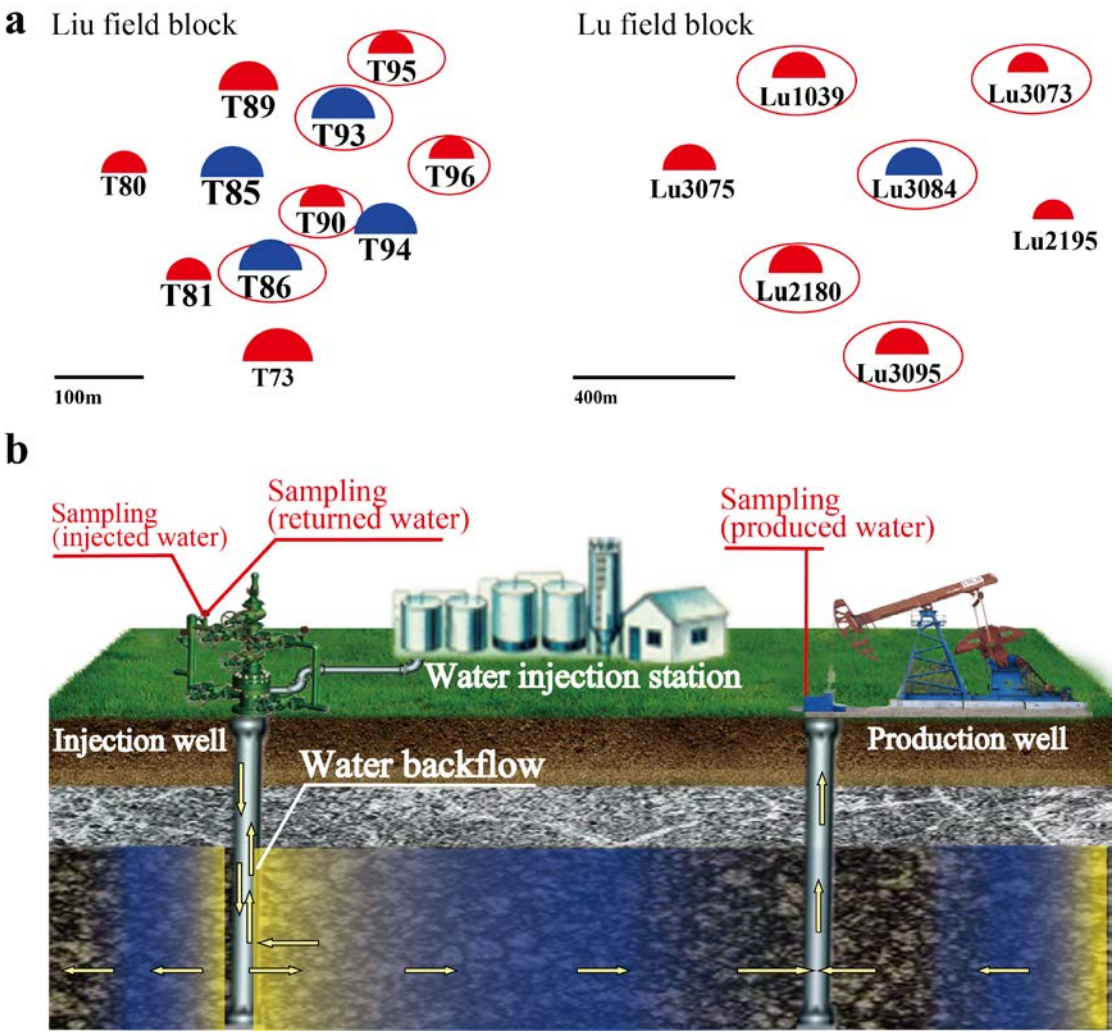

a
Liu field block
T95
T89
T93
T96
T80
T85
T90
T94
T81
T86
T73
100m
Lu field block
Lu1039
Lu3073
Lu3075
Lu3084
Lu2195
Lu2180
Lu3095
400m
b
Sampling
(injected water)
Sampling
(returned water)
Sampling
(produced water)
Water injection station
Injection well
Water backflow
Production well

Fig. 1

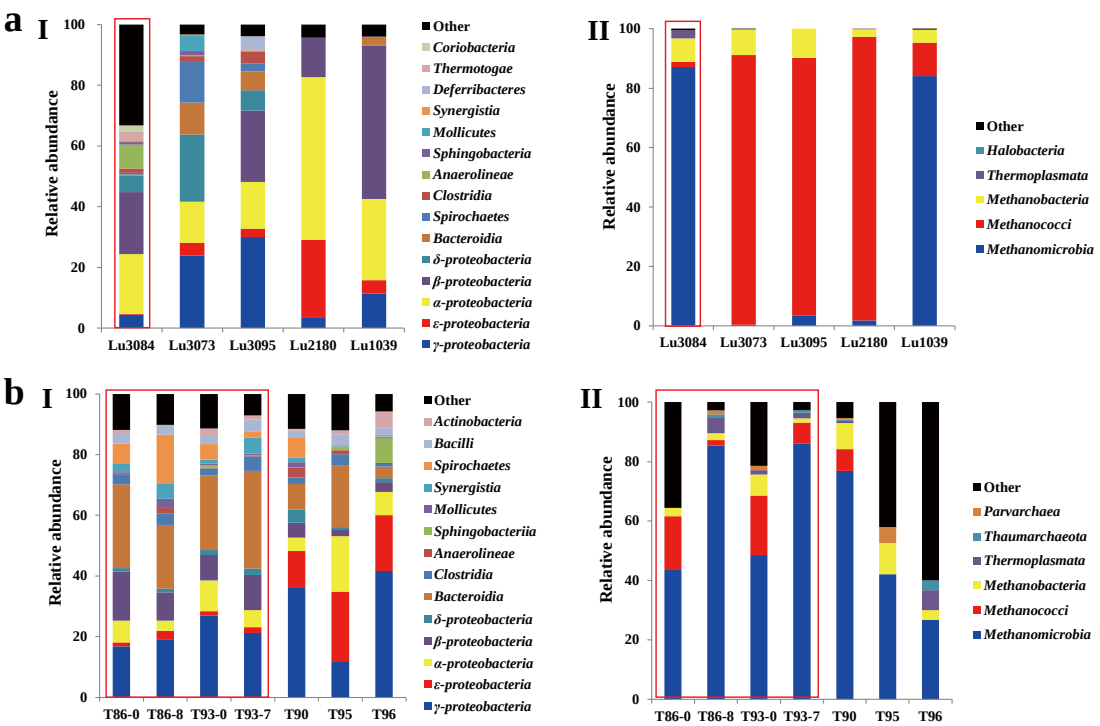
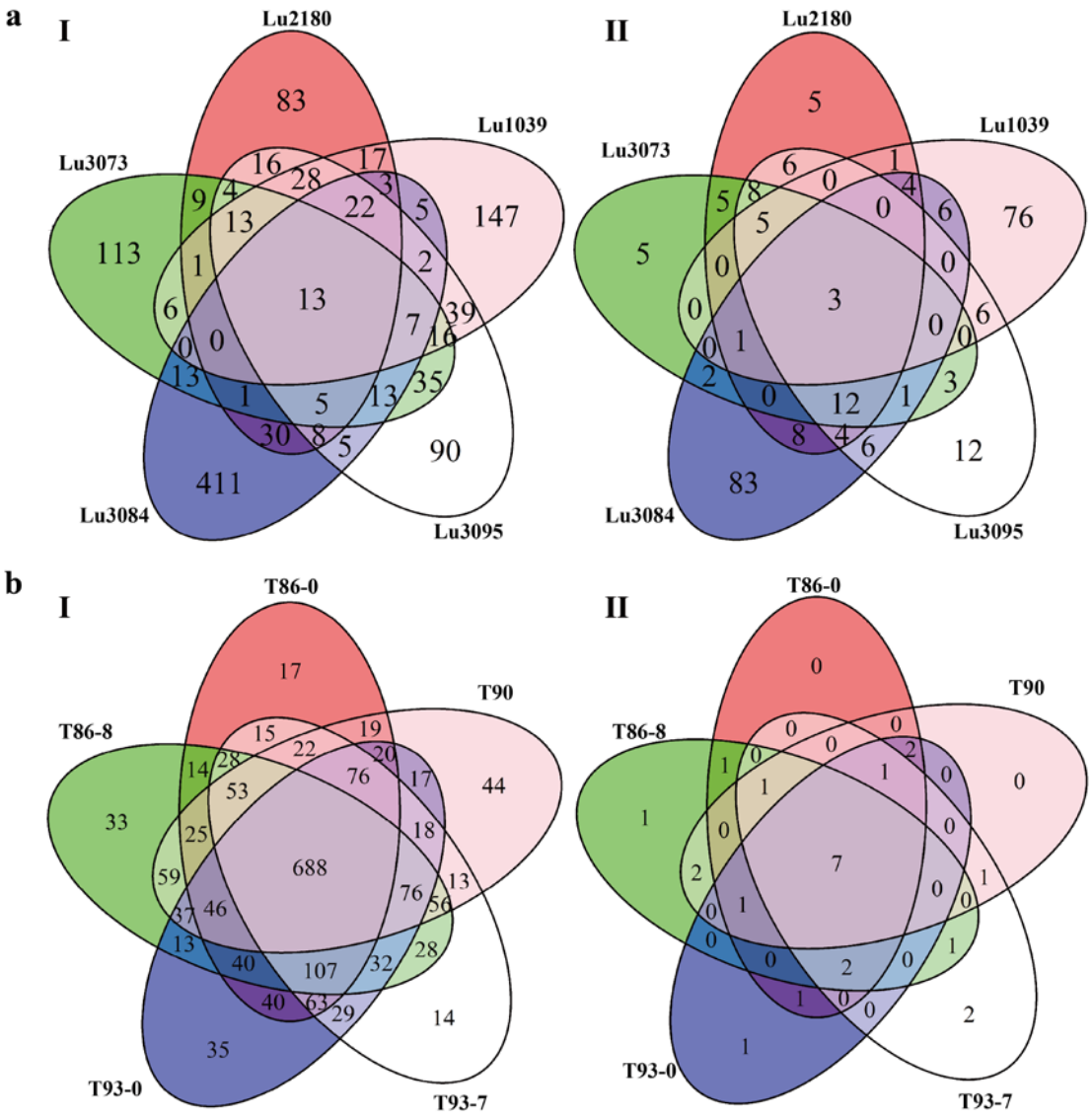

a
I
II
Relative abundance
100
80
60
40
20
0
Lu3084
Lu3073
Lu3095
Lu2180
Lu1039
Other
Coriobacteria
Thermotogae
Deferribacteres
Synergistia
Mollicutes
Sphingobacteria
Anaerolineae
Clostridia
Spirochaetes
Bacteroidia
δ-proteobacteria
β-proteobacteria
α-proteobacteria
ε-proteobacteria
γ-proteobacteria
Halobacteria
Thermoplasmata
Methanobacteria
Methanococci
Methanomicrobia
b
T86-0
T86-8
T93-0
T93-7
T90
T95
T96
Actinobacteria
Bacilli
Sphingobacteriia
Parvarchaea
Thaumarchaeota

Fig. 2



654

655 **Fig. 3**

656

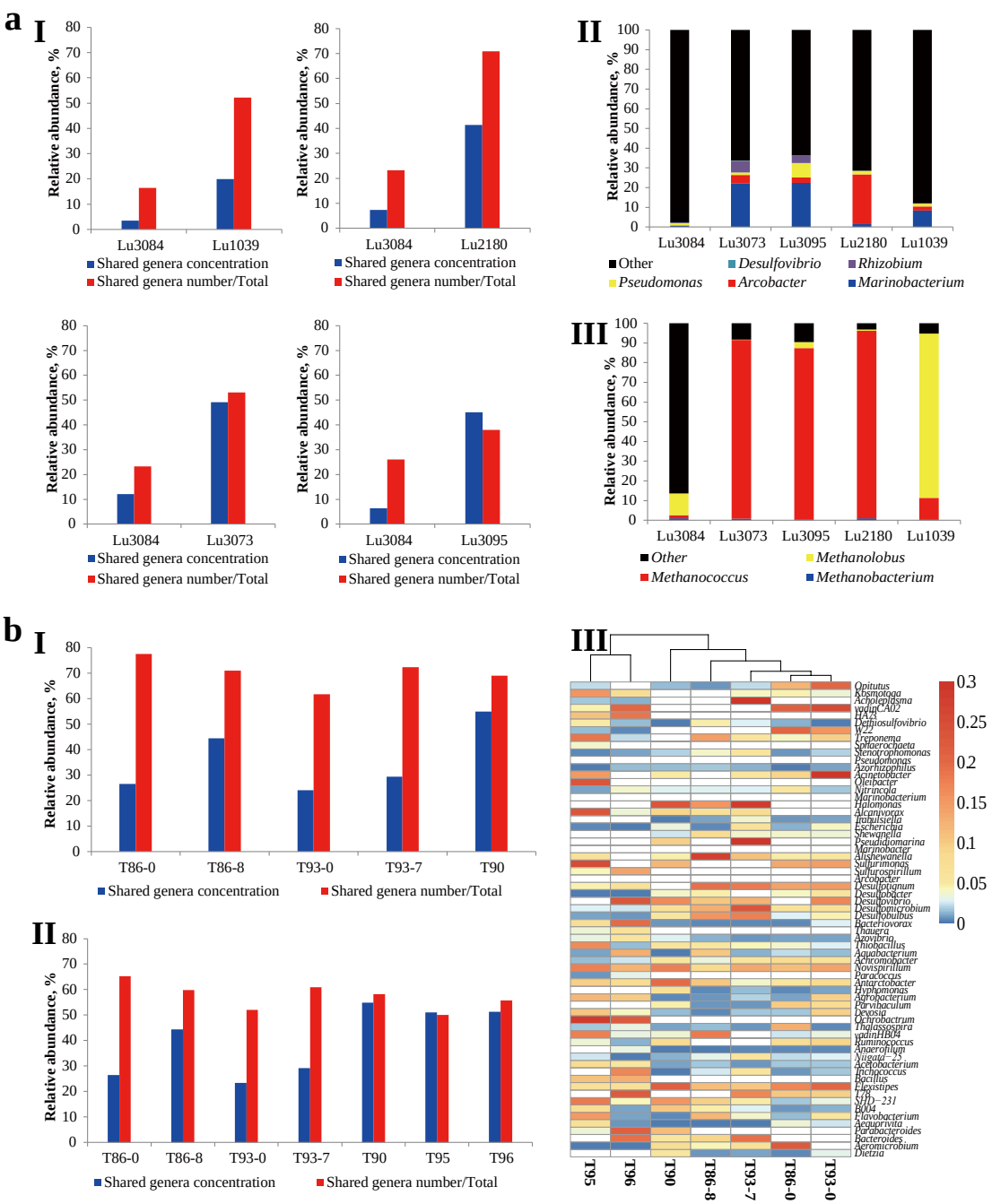
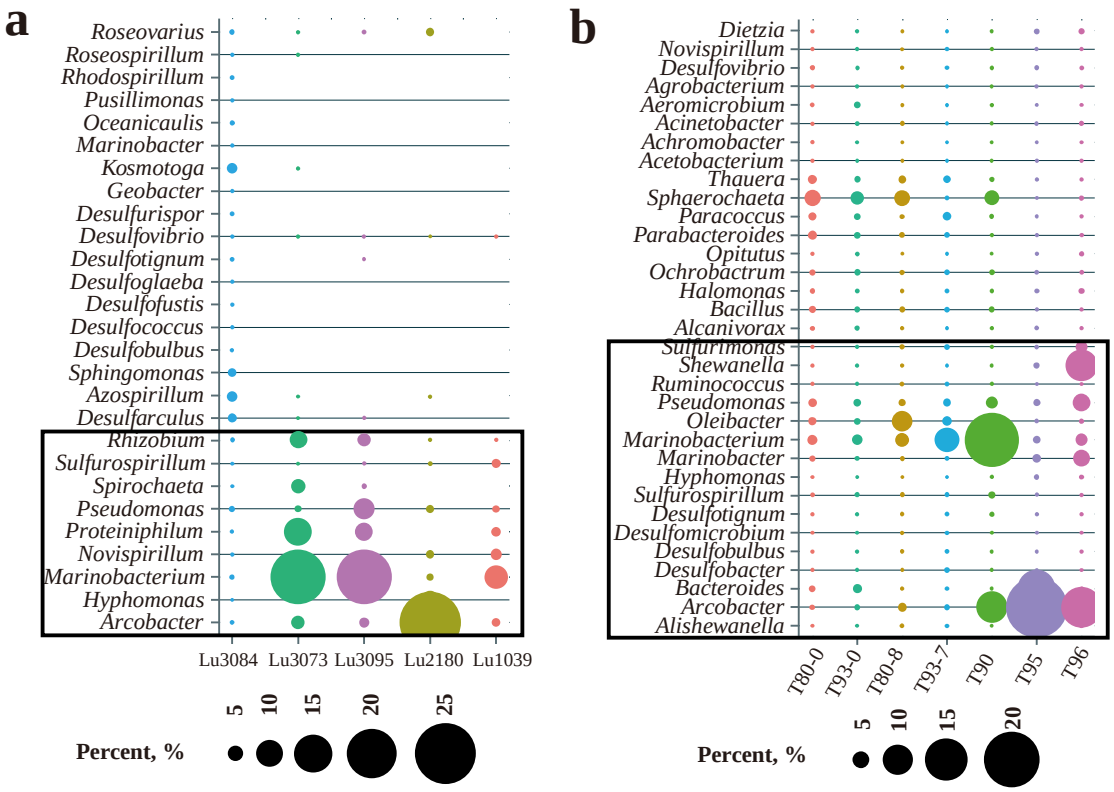

a
I
Relative abundance, %
Lu3084
Lu1039
Lu2180
Lu3073
Lu3095
Shared genera concentration
Shared genera number/Total
II
Other
Desulfovibrio
Rhizobium
Pseudomonas
Arcobacter
Marinobacterium
III
Methanococcus
Methanolobus
b
T86-0
T86-8
T93-0
T93-7
T90
T95
T96
0.3
0.25
0.2
0.15
0.1
0.05
0

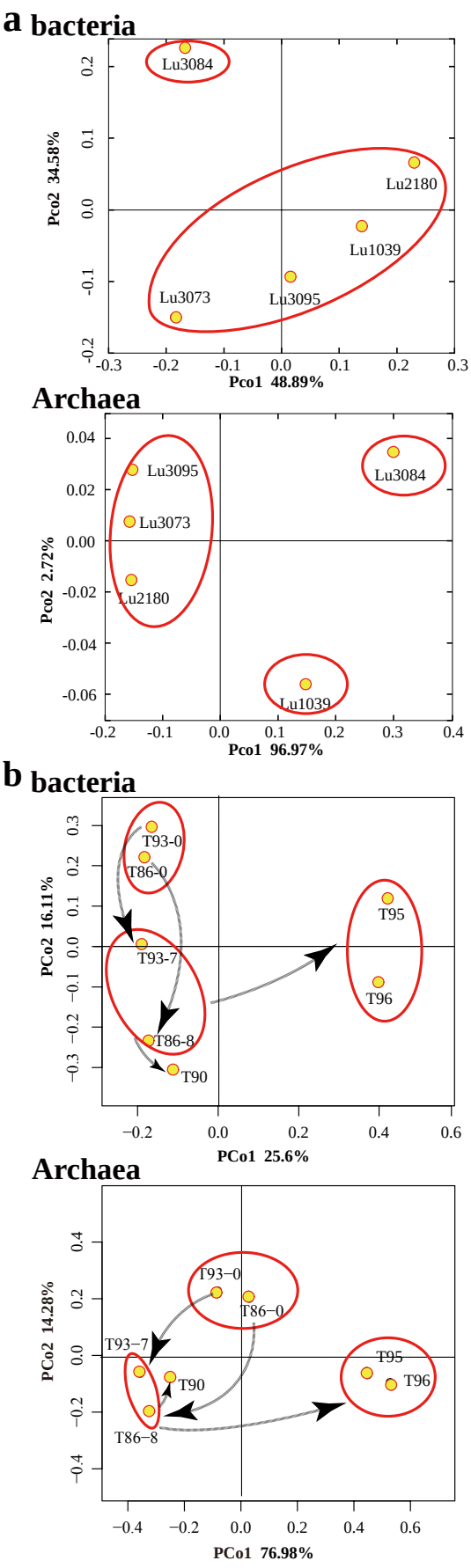
Fig. 4



660

661 **Fig. 5**

662


a bacteria
Lu3084
Lu2180
Lu1039
Lu3073
Lu3095
Pco2 34.58%
Pco1 48.89%
Archaea
Lu3095
Lu3084
Lu3073
Lu2180
Lu1039
Pco2 2.72%
Pco1 96.97%
b bacteria
T93-0
T86-0
T95
T93-7
T96
T86-8
T90
PCo2 16.11%
PCo1 25.6%
Archaea
T93−0
T86−0
T93−7
T95
T96
T90
T86−8
PCo2 14.28%
PCo1 76.98%

663

664 **Fig. 6**