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. The microbial community composition of water-flooding petroleum reservoirs is of great interest because it is strongly related to the enhancement of oil recovery. However, our knowledge about the relationship between microbial communities in injection and production wells is still very limited. The present study investigated the differences in microbial communities in 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. The results indicate that a small number of microbial populations are shared between the injected and produced water samples in the sandstone reservoir, whereas a large number of microbial populations are shared in the conglomerate reservoir. Consistently, the community structure exhibited large differences between the injected and produced water samples, with the shared populations accounting for a minor fraction of the injected water, but dominating the produced water in both reservoirs. This suggests that the community is reassembled as the injected water flows into the production wells. The results imply that microbial communities have significant differences between injection and production wells, in particular, the community composition and the relative abundance, which have a close relationship with the sieve effect of strata and the dissolved oxygen.

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

1 **Introduction**

Water-flooding is an efficient oil recovery process that is employed worldwide. After long-term water-flooding, diverse microbial populations inhabit the petroleum reservoirs. These microorganisms and their metabolites, such as polysaccharide and biosurfactants, 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). These characteristics have been used to improve oil recovery. With an increasing global energy demand and the depletion of oil reserves, microbial enhanced oil recovery (MEOR) is currently studied extensively. To date, a large number of laboratory studies and field trials have been performed on stimulating reservoir microorganisms to improve oil recovery (Abdel-Waly, 1999;Zhang et al., 2012;Bao et al., 2009;Gao et al., 2013;Li et al., 2014).

Microbial populations are important components of reservoir ecosystems, and play critical roles in the microbial enhancing of the oil recovery process. Recently, culture-dependent and -independent methods, in particular, 16S rRNA-based molecular identification methods, have revealed diverse microorganisms inhabiting petroleum reservoirs (Al-Bahry et al., 2013;Kumaraswamy et al., 2011;Lenchi et al., 2013;Okoro et al., 2014). However, the relationship between microbial communities in injection and production wells remains poorly understood. Based on the 16S rRNA gene clone library method, several studies 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). Unfortunately, these studies did not compare the differences of microbial community composition between injection and production water samples. Furthermore, 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 the 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 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. The results presented here expand our knowledge on the relationships of microbial communities between injection and production water samples of water-flooding petroleum reservoirs.

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 selected 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 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). Although the injection wells have a direct influence on the neighboring production wells, the conglomerate reservoir heterogeneity is very strong.

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 1. 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 among 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.

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 were collected from the neighboring production wells 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) 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. Following immediate transportation to

142 the laboratory, the residual oil was first removed by heating the sample to 60°C for 30
143 min and by phase separation in sterilized separatory funnels.
144
145 Microbial cells were then collected from 5 L of each water sample by centrifugation
146 at 4°C for 15 min at 10,000 × g in a high-speed centrifuge (Beckman, CA 92821,
147 USA). The cell deposits collected from the same sampling location were mixed and
148 resuspended with TE buffer (Tris 80 mM, EDTA 40 mM, pH 8.0), and then lysed
149 using a mini bead-beater (BioSpec, Bartlesville, OK 74005, USA) at 200 rpm for 1
150 min at room temperature with 0.1 mm glass beads. After bead beating, lysozyme was
151 added (final concentration of 1 mg/ml), and the samples were incubated at 37°C for 1
152 h. Following the lysozyme treatment, 120 μL sodium-dodecyl sulphate (20% SDS,
153 W/V) was added and the samples were incubated at 65°C for 60 min. Total genomic
154 DNA was then extracted from the suspension solution using an AxyPrepTM Genomic
155 DNA miniprep kit (Axygen Biosciences, Tewksbury, MA 01876, USA) according to
156 the manufacturer's instructions and stored at -80°C for subsequent study.
157
158 **2.3 Pyrosequencing of partial 16S rRNA genes and sequence analysis**
159 Broadly conserved primers 27F (5'-AGA GTT TGA TCC TGG CTC AG-3') and
160 533R (5'-TTA CCG CGG CTG CTG GCA C-3') were used to amplify the bacterial
161 16S rRNA gene, while primers 344F (5'-ACG GGG YGC AGC AGG CGC GA-3')
162 and 915R (5'-GTG CTC CCC CGC CAA TTC CT-3') were used to amplify the
163 archaeal 16S rRNA gene. PCR reactions were performed following the protocol
164 described in the Supporting Information (SI). Replicate PCR products of the same
165 sample were mixed in a PCR tube. Then, the amplicons from each reaction mixture
166 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 communities 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. The individual rarefaction, Shannon, and Phylogenetic diversity curves tended to approach the saturation plateau (Fig. S1). Based on the results of qPCR, the copy number of bacterial 16S rRNA ranged from 1.5×10^{6} to 8.25×10^{6} copies ml^{-1}, while archaeal 16S rRNA ranged from 8.5×10^{3} to 5.75×10^{4} copies ml^{-1} in the water samples. 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).

In the sandstone reservoir, 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%). *Methanomicrobia* accounted for 84.03% in the produced water at 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* produce CH_4 when growing with methylamine as carbon source, while *Methanococcus* use H_2 and formate as carbon sources.

3.2 Microbial communities in the conglomerate reservoir

Between 52719 to 129106 16S rRNA gene sequences were analyzed and assigned to 2623 to 3414 genus-level OTUs. 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). The copy number of 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 both the injected and produced water, composing 85.7%–94.1% of all bacterial communities. Similar to the sandstone reservoir, more *Proteobacteria* were detected in the produced water samples. At the class level, *Gammaproteobacteria*, *Epsilonproteobacteria*, *Alphaproteobacteria*, *Betaproteobacteria*,

Deltaproteobacteria, *Bacteroidia*, *Bacilli*, and *Clostridia* composed 74.5%–83.7% of the bacterial communities in both the injected and produced water samples (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 in both the injected and produced water, and composed 64.3%–94.6% of the archaeal communities. 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 injection and production wells

The shared microbial OTUs and genera between communities in the injected and produced water samples 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 injected water and each produced water sample in the sandstone reservoir (Fig. 3a). These shared bacterial OTUs accounted for 4.6–24.71% of the total bacterial community in the injected water, and 43.23–76.18% in each produced water sample (Fig. 4a I). Furthermore, only 13 bacterial and 3 archaeal OTUs were shared by both the injected and produced water samples (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 detected in the injected and produced water samples, comprising 13.58% of the total archaea in the injected water, and 90.4%–96.89% in each of the produced waters (Fig. 4a III).

In the conglomerate reservoir, most of the OTUs and genera were simultaneously detected in the injected and produced water samples (Fig. 3b and 4b). 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 injection and production wells

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 injected water, while *Arcobacter*, *Marinobacterium*, *Pseudomonas*, *Hyphomonas*, *Novispirillum*, *Proteiniphilum*, *Spirochaeta*, and *Rhizobium* were highly abundant in the produced water. In the conglomerate reservoir, higher amounts of *Paracoccus*, *Bacillus*, *Ochrobactrum*, *Parabacteroides*, *Sphaerochaeta*, *Thauera*, *Halomonas*, and *Alcanivorax* were detected in the injected water, 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*,

317 *Halomonas*, *Spirochaeta*, *Bacillus*, *Thauera*, *Halomonas*, and *Bacteroides* are
318 microaerophilic bacteria, facultative anaerobes, or anaerobes.
319
320 To further investigate the microbial distribution in injection and production wells,
321 hierarchical clustering and Unifrac PCoA were performed based on microbial OTUs
322 abundance and phylogenetic relationships. In the sandstone reservoir, hierarchical
323 clustering showed that the community in the injected water was distinct from that of
324 the produced water (Fig. S3). Weighted PCoA distinguished the bacterial community
325 of the injected water from that of the production wells, while communities of the
326 production wells were placed at a comparatively decentralized position (Fig. 6a I).
327 Similar to the bacterial communities, hierarchical clustering and PCoA distinguished
328 the archaeal community of the injected water from those of the production wells,
329 whereas production wells were placed at a close proximity (Fig. 6a II). In the
330 conglomerate reservoir, communities of water samples collected from the wellhead of
331 injection wells clustered into a group in the PCoA plot, indicating that communities
332 remained unchanged before injected water flowed into the injection wells (Fig. 6b).
333 Communities in the water samples collected from the downhole of injection wells and
334 neighboring production well T90 clustered into one group, while production well T95
335 and T96 clustered into another (Fig. 6b). This shows that the microbial community
336 reassembled during the process of the injected water flowing into the reservoir strata
337 and each production well.
338
339 **4 Discussion**
340 MEOR technique is generally classified into exogenous microbial flooding and
341 indigenous microbial flooding. The former includes injection of exogenous

342 microorganisms and injection of ex-situ produced products into reservoirs to enhance
343 oil recovery (Zobell, 1947). This is an effective way to quickly improve oil recovery.
344 However, because of the sieve effect of strata on microbial cells, the injected
345 microorganisms are generally difficult to migrate into reservoir strata (Youssef et al.,
346 2009;Brown, 2010). Indigenous microbial flooding technique improves oil recovery
347 by introducing oxygen and salts through water-based injection to stimulate reservoir
348 microorganisms (Belyaev SS et al., 1998). Despite the validity in field trial, this
349 technology also has some limitations, in particular, instability during microbial
350 flooding process. Microbial community diversification has been found to have a
351 significant influence on oil displacement efficiency (Li et al., 2014). Because that
352 reservoir microbial populations and their metabolites play an important role in the
353 enhancement of oil recovery, microbial community composition and distribution in
354 physically and geochemically diverse reservoirs has been extensive studied (Al-Bahry
355 et al., 2013;Kumaraswamy et al., 2011;Lenchi et al., 2013;Okoro et al., 2014).
356 However, the relationship between microbial communities in injection and production
357 wells remains poorly understood. We have therefore compared the differences of
358 microbial community composition between injection and production water samples,
359 and observed the microbial community diversification and succession as the injected
360 water flows into the production wells.
361
362 Molecular methods have been widely used to assess the microbial diversity of
363 petroleum reservoirs. Compared to the traditional 16S rRNA gene clone library and
364 sequencing, high-throughput sequencing has generated hundreds of thousands of short
365 sequences, and significantly improved our ability to compare microbial populations
366 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 depth was approximately 10–20 folds of pyrosequencing used in the sandstone reservoir, and 50–400 folds of the 16S rRNA gene clone library (assuming 300 clones per library). However, the current sequencing depth is still limited for detecting archaeal populations. As a result, 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, suggesting the need for deeper sequencing for the detection of rare archaeal populations using this sequencing method. 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 and Shannon curves tended to approach the saturation plateau, suggesting that this sequencing depth was enough for the detection of major bacterial and archaeal communities.

The community structure exhibited large differences between the injected and produced water samples. Differences in microbial communities may result from a number of different factors. The niche-based processes are supposed to be the primary drivers for the community diversification, and environmental variables such as salinity, pH, nitrogen, and phosphorus identified as the major determinants of microbial community composition (Kuang et al., 2013). However, few differences in cations and anions among the injected and produced water samples were observed. Petroleum reservoirs represent extreme anaerobic environments with multiphase fluids of oil, gas and water. Therefore, the subtle differences in the reservoir strata, in particular, the permeability, porosity, and dissolved oxygen, may exert a significant

influence on the microbial 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 that most microbial populations detected in the injected water were not detected in the production wells. It is reasonable to speculate that the abundant microorganisms in the injected water do not reach the production wells in this sandstone reservoir. Based on previous research, the main reason for this 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 the reservoir strata. In contrast, we found that almost all OTUs and genera detected in the injected water were also observed downhole of the injection and neighboring production wells in the heterogeneous reservoir, which has a similar permeability but shorter inter-well spacing, compared with the sandstone reservoir. It appears that most microbial populations in the injected water migrated into the oil-bearing strata and reached the production wells in the conglomerate reservoir. 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. To solve this issue, labelled strains, such as ones containing green fluorescent protein, may be a suitable way to investigate microbial migration in petroleum reservoirs.

Compared with the sandstone reservoir, a large number of microbial populations were simultaneously detected in the injected and produced water samples in the conglomerate reservoir. However, the shared OTUs and genera accounted for a minor fraction of the injected water in both reservoirs, whereas they dominated the produced water in both reservoirs, suggesting that the microbial community was reassembled as the injected water flowed into the production wells. Dissolved oxygen, which is known to be strongly related to microbial growth and metabolism (Gao et al., 2013), may be the main factor influencing the community structures. Although in situ oxygen concentrations were not recorded in this study, more aerobic bacteria, including *Sphingomonas*, *Azospirillum*, *Paracoccus*, *Ochrobactrum*, *Alcanivorax*, and *Hydrogenophilaceae* were detected in the injected water, while microaerophilic bacteria, facultative anaerobes, and anaerobes, including *Pseudomonas*, *Rhizobium*, *Arcobacter*, *Halomonas*, *Spirochaeta*, and *Bacteroides*, were found to have higher relative abundance in the produced water (Fig. 5). Apart from the dissolved oxygen, another striking factor influencing microbial distribution in the injected water and the production wells may have been the crude oil, in particular, the saturates and aromatic components. 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 downhole of injection and production wells.

This study compared the differences in microbial community composition between injection and production water samples 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.

This study investigated the relationship shared by microbial communities in injection and production water samples, and found the significant differences between microbial communities in the injection and production water samples. However, it is less rigorous to make a conclusion on the transport of microbial populations in the reservoir strata by the current results. To solve the problem, injecting labelled strains containing green fluorescent protein into reservoirs may bring novel insight and greater predictive power to investigate microbial migration in reservoir strata. Therefore, the 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, we comprehensively surveyed the relationship shared by microbial communities in injection and production water samples in homogeneous sandstone and heterogeneous conglomerate reservoirs. The results suggest that the microbial communities have significant differences between the injection and production water samples. Even if most microbial populations were shared, the relative abundance of shared populations exhibited large differences between the injected and produced water samples. Water backflow in the injection wells suggested that the microbial community was reassembled during the process of the injected water flowing into the production wells.

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.

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.

Kumaraswamy, R., Ebert, S., Gray, M. R., Fedorak, P. M., and Foght, J. M.: Molecular- and cultivation-based analyses of microbial communities in oil field water and in microcosms amended

with nitrate to control H2S production, Appl. Microbiol. Biotechnol., 89, 2027–2038, 2011.

Lenchi, N., İnceoğlu, Ö., Kebbouche-Gana, S., Gana, M. L., Llirós, M., Servais, P., and García-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, e66588, 2013.

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. X., Dong, H. P., and Ma, T.: Microbial Abundance and Community Composition Influence Production Performance in a Low-Temperature Petroleum Reservoir, Environ Sci Technol, 48, 5336-5344, 2014.

Magoc, T., and Salzberg, S. L.: FLASH: fast length adjustment of short reads to improve genome assemblies, Bioinformatics, 27, 2957-2963, 2011.

Okoro, C., Smith, S., Chiejina, L., Lumactud, R., An, D., Park, H. S., Voordouw, J., Lomans, B. P., and Voordouw, G.: Comparison of microbial communities involved in souring and corrosion in offshore and onshore oil production facilities in Nigeria, J Ind Microbiol Biotechnol, 41, 665-678, 2014.

Ren, H. Y., Zhang, X. J., Song, Z. Y., Rupert, W., Gao, G. J., Guo, S. X., and Zhao, L. P.: Comparison of microbial community compositions of injection and production well samples in a long-term water-flooded petroleum reservoir, Plos One, 68, e23258, 2011.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., Sahl, J. W., Stres, B., Thallinger, G. G., Van Horn, D. J., and Weber, C. F.: Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities, Appl. Environ. Microbiol., 75, 7537–7541, 2009.

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, e33535, 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 Biot, 98, 4209-4221, 2014.

Youssef, N., Elshahed, M. S., and McInerney, M. J.: Microbial processes in oil fields: culprits, problems, and opportunities, Adv. Appl. Microbiol., 66, 141–251, 2009.

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 of an indigenous microbial enhanced oil recovery field trial on microbial community structure in a high pour-point oil reservoir, Appl. Microbiol. Biotechnol., 95, 811–821, 2012.

Zobell, C.: Bacterial release of oil from sedimentary materials, Oil Gas Journal, 46, 62-65, 1947.

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. T86, T93 and Lu3084 are the selected injection wells, and T90, T95, T96, Lu1039, Lu2180, Lu3073 and Lu3095 the production wells.

Fig. 2. Relative proportion of microbial taxa at class level in the injection and production 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.

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.

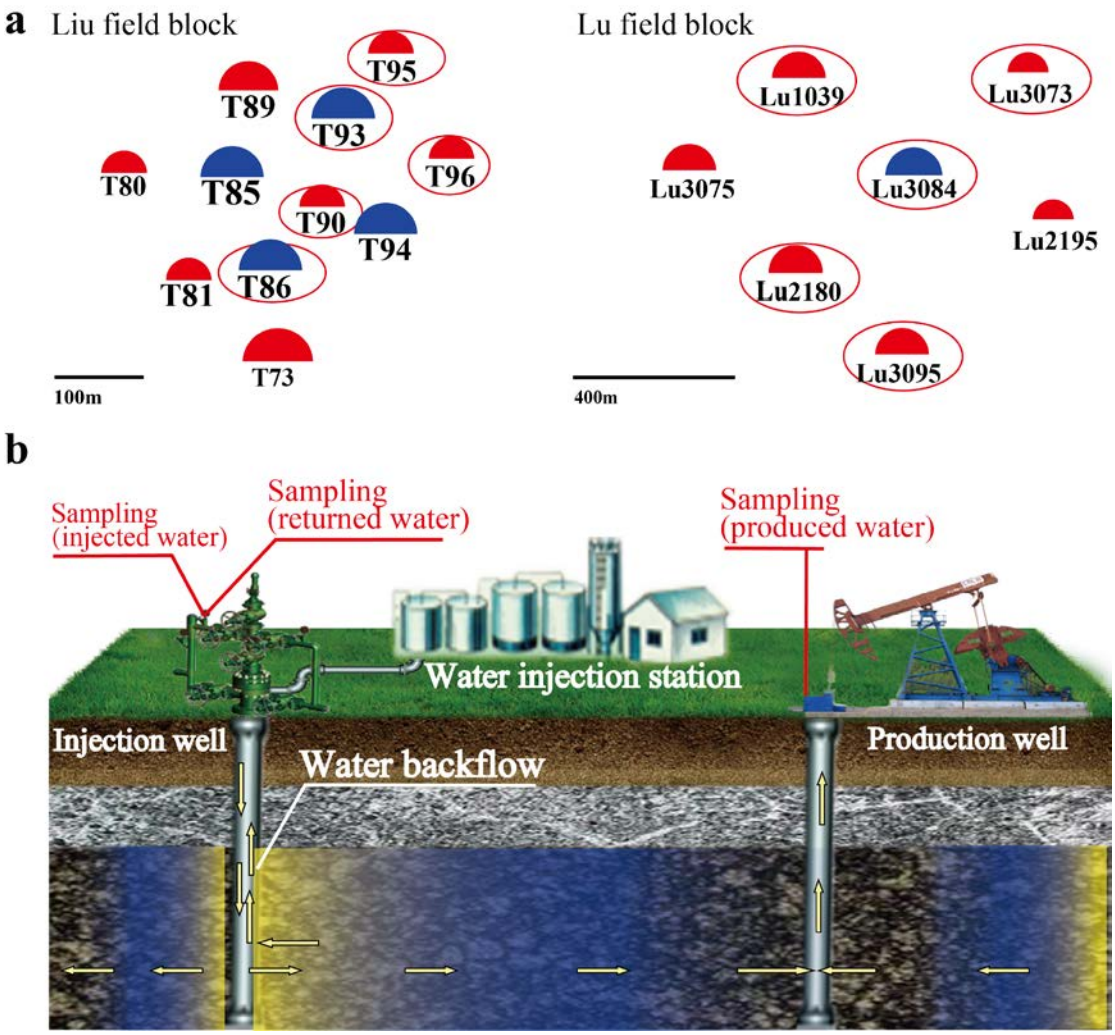
Fig. 6. Principal coordinate analysis of microbial communities. (a) Sandstone and (b)

conglomerate reservoirs. I: bacterial community distribution; II: 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.

595 **Table 1** Characterization of the reservoir characteristics and the collected water samples

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						
Well number	**Lu3084**	**Lu1039**	**Lu2180**	**Lu3073**	**Lu3095**	**T86–0**	**T86–8**	**T93–0**	**T93–7**	**T90**	**T95**	**T96**
Well type	Injection well	Production well				Injection well				Production well		
Mineralization	10850	11690	11170	10545	11102	10101	11313	11399	13991	13203	8997	9710
Total nitrogen, mg/L	15.1	11.5	10.6	12.7	11.6	7.6	6.8	10.2	11.5	5.7	6.3	8.5
Total phosphorus, mg/L	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^+	4524.9	4803.1	4565.3	4308.7	4486.9	3364.3	3630.1	3801.6	4348.5	4014.2	3097	3139.8
Mg^{2+}	21.7	32.07	31.55	26.03	28.83	33.1	63.05	28.37	63.05	68.09	17.46	50.2

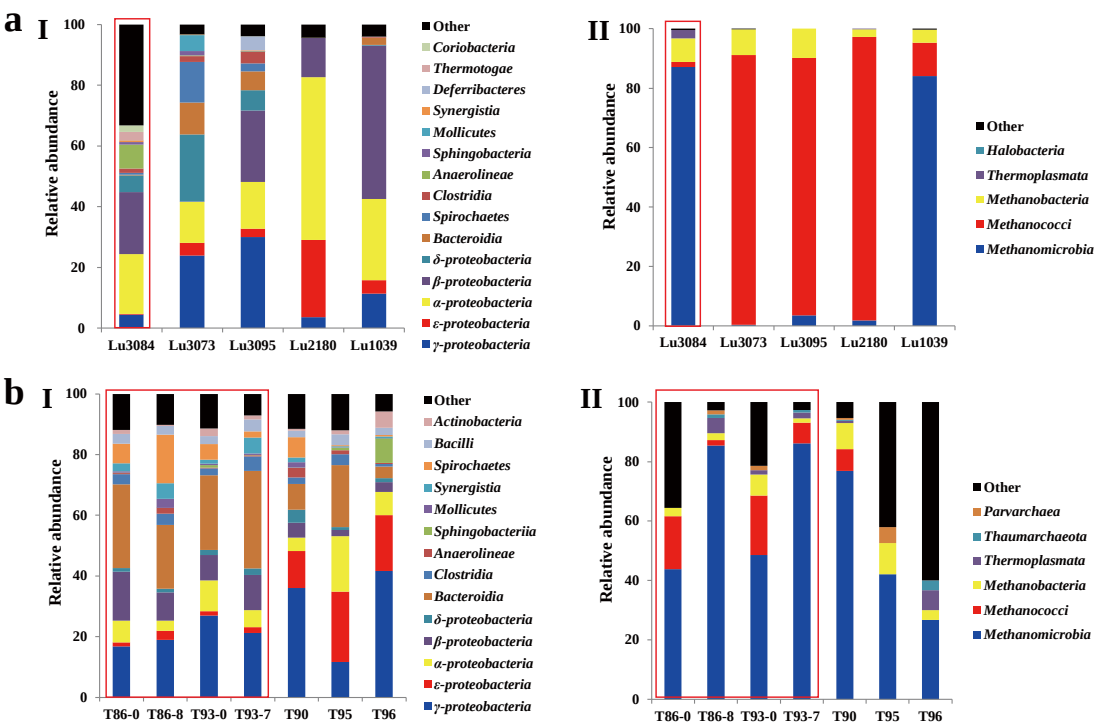
Ca^{2+}	191.3	281.9	284.7	181.6	216.4	70.18	77.98	72.78	77.98	96.17	86.36	108.66
Cl^{-}	5640	6125	5820	5160	5850	3010.24	3629.99	2921.7	3452.92	3098.77	3816.21	3405.92
SO_4^{2-}	116.2	14.04	4.86	23.13	8.86	483.4	89.4	523	362.4	89.9	65.6	165.1
HCO_3^{-}	356	434.1	464	846.29	511.4	3140.1	3823	4051.8	5686.9	5836.6	1914.9	2840.6



596

597 **Fig. 1**

598



599

600 **Fig. 2**

601

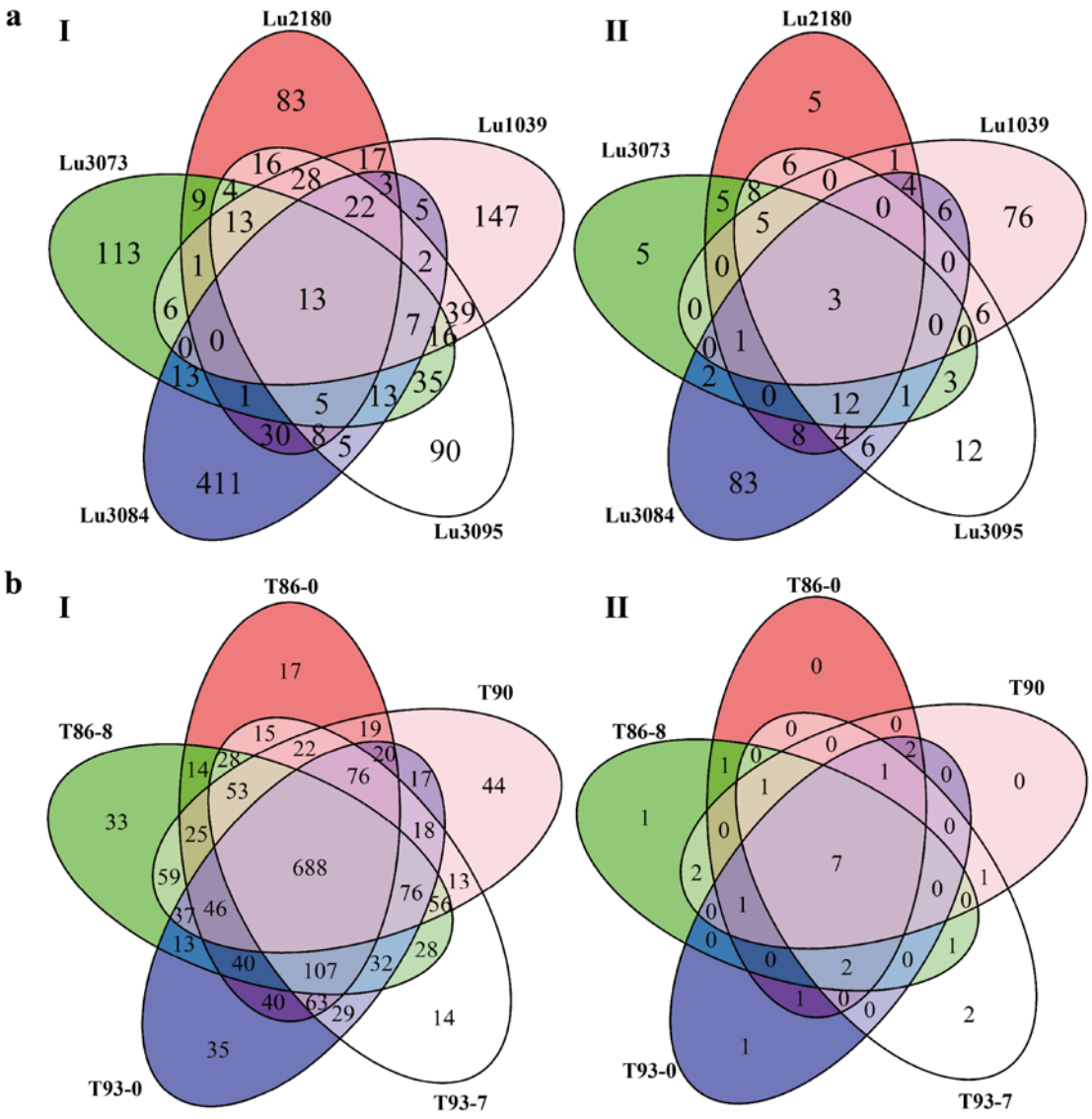

a
I
Lu2180
Lu1039
Lu3073
Lu3084
Lu3095
83
147
113
13
411
90
II
Lu2180
Lu1039
Lu3073
Lu3084
Lu3095
5
76
5
3
83
12
b
I
T86-0
T90
T86-8
T93-0
T93-7
17
44
33
688
35
14
II
T86-0
T90
T86-8
T93-0
T93-7
0
0
1
7
1
2

Fig. 3

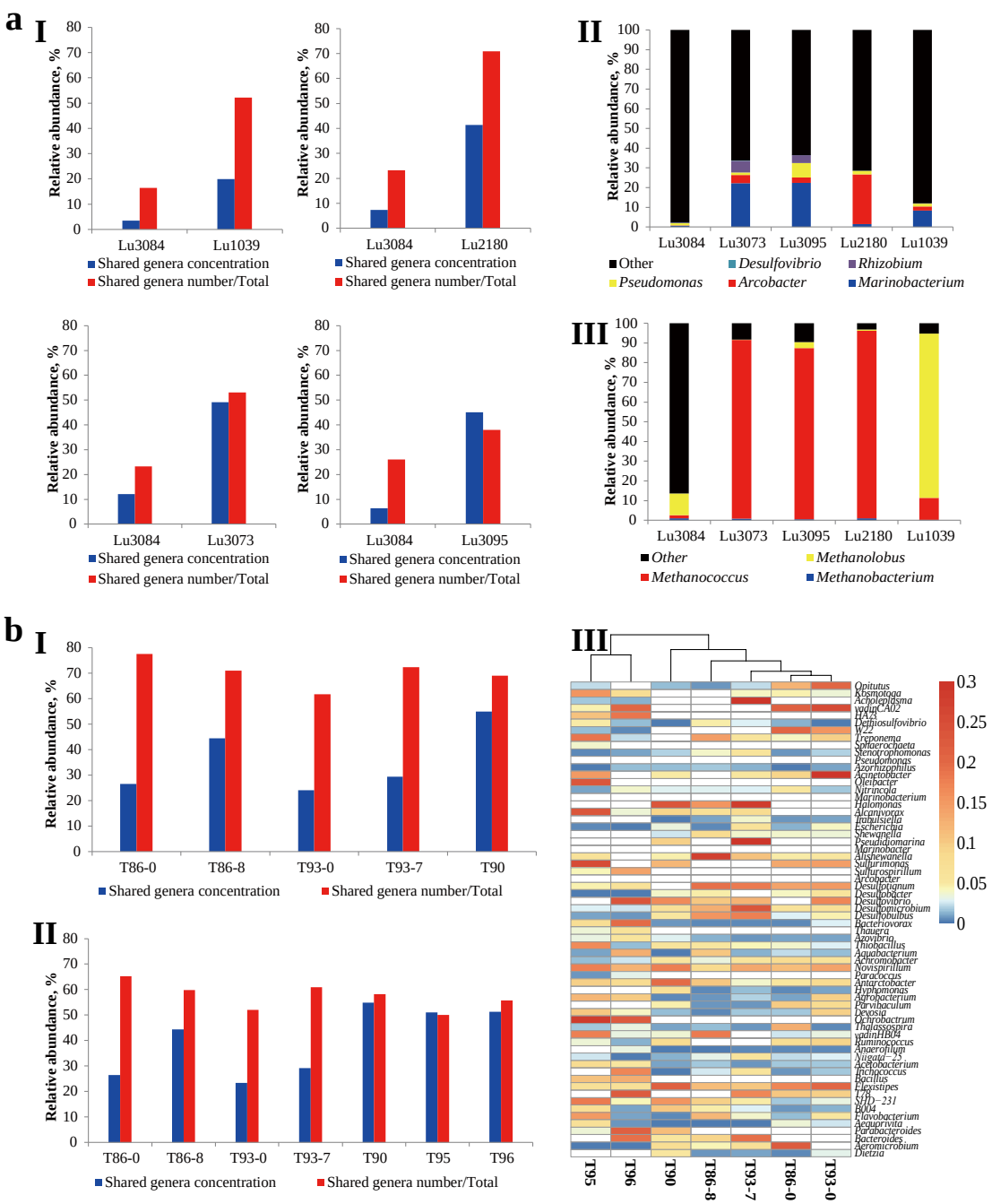
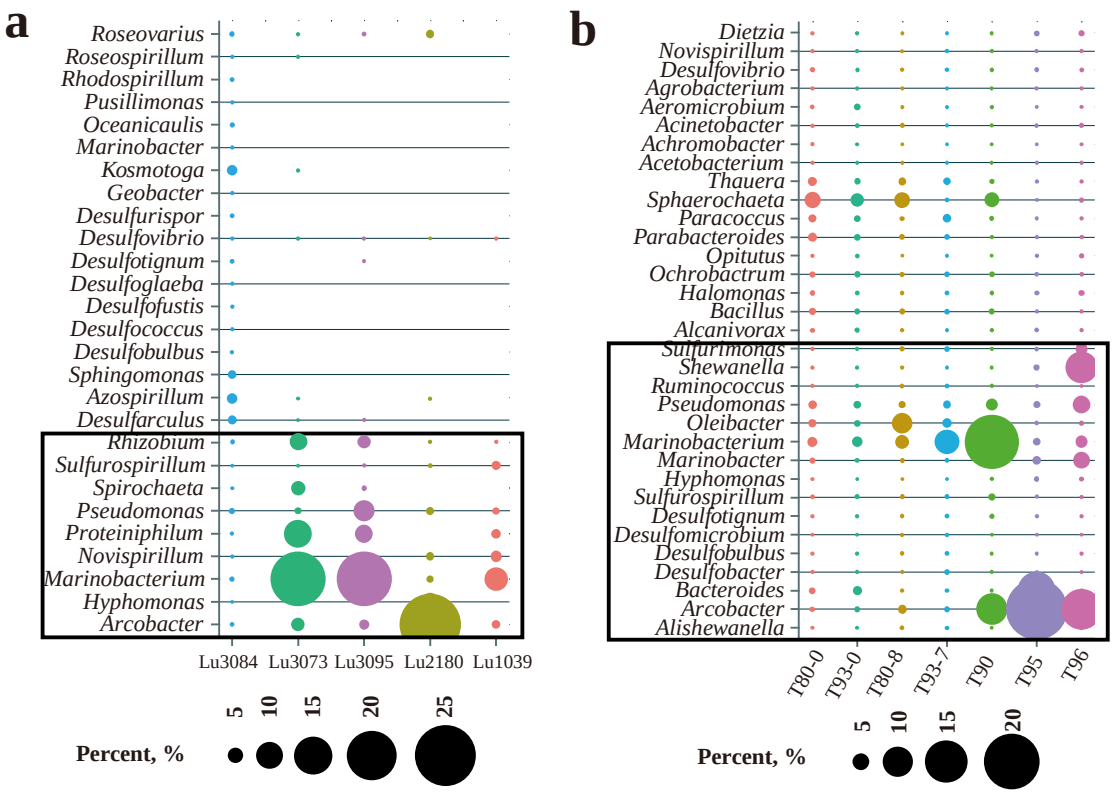

a
I
Relative abundance, %
Lu3084
Lu1039
Lu2180
Lu3073
Lu3095
Shared genera concentration
Shared genera number/Total
II
Other
Desulfovibrio
Rhizobium
Pseudomonas
Arcobacter
Marinobacterium
III
Methanococcus
Methanolobus
b
T86-0
T86-8
T93-0
T93-7
T90
T95
T96
0.3
0.25
0.2
0.15
0.1
0.05
0

Fig. 4



608

609 **Fig. 5**

610

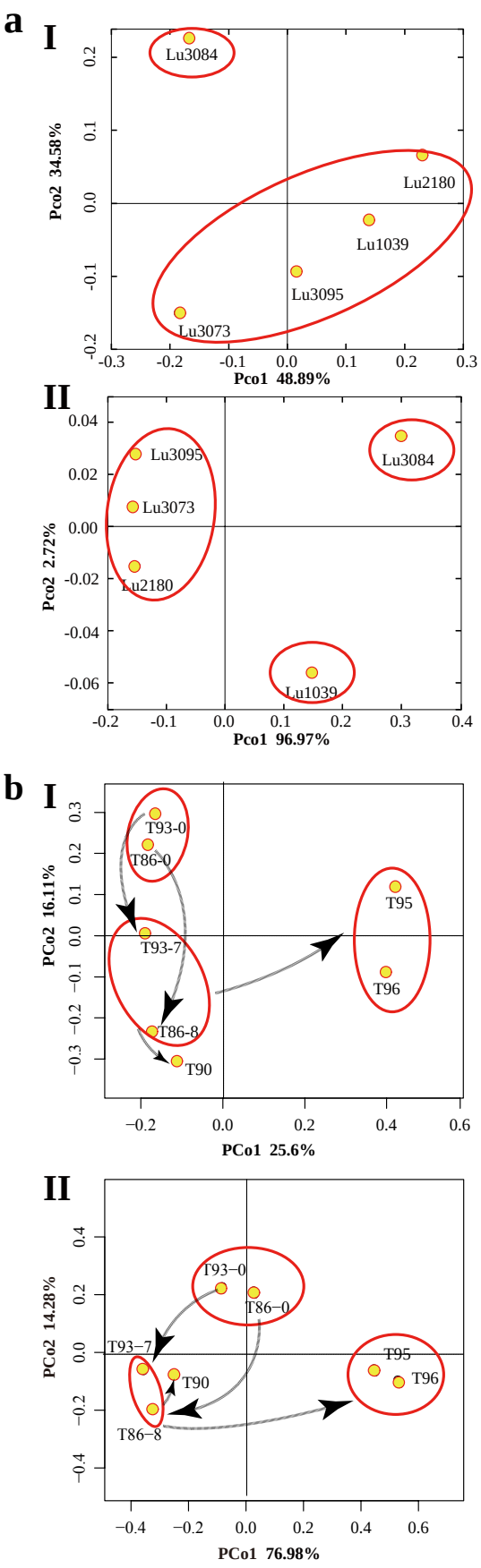

a
I
Lu3084
Lu2180
Lu1039
Lu3095
Lu3073
Pco2 34.58%
Pco1 48.89%
II
Lu3095
Lu3073
Lu2180
Lu3084
Lu1039
Pco2 2.72%
Pco1 96.97%
b
I
T93-0
T86-0
T93-7
T86-8
T90
T95
T96
PCo2 16.11%
PCo1 25.6%
II
T93−0
T86−0
T93−7
T90
T86−8
T95
T96
PCo2 14.28%
PCo1 76.98%

Fig. 6