Dear editor Jia

We would like to thank you and the reviewers for the time and efforts to review our manuscript. The constructive comments and suggestions were considered carefully, and we have addressed these issues raised point by point. We hope that our revised manuscript will be considered suitable for publication in *Biogeosciences*.

Thank you for your time and efforts.

Sincerely yours,

Dr. Ting Ma

Authors' responses to editor Jia

Initial Decision: Reconsider after major revisions (12 Apr. 2015)

Comments to the Author:

Your manuscript has been reviewed by two external reviewers with relevant expertise and your responses to the comments are adequate. New experiments and data mining were also conducted, and your new figure 1 showing the experimental flow diagram is particularly welcomed. However, there are still major concerns that preclude publication of your manuscript at its current form.

Specific comments:

(1) First of all, please provide all information about pyrosequencing. Please *make a supplementary table showing* ***the number of raw sequence, read length, high-quality sequence number*** *and so on*. In addition, I strongly suggest that ***each of sequence data (individual location) be assigned with an unique accession number***. In other words, please submit your high-throughput sequencing data to the GenBank on the basis of individual sample tested in this study, but not simply as a whole with data pooled from all samples. This will be of great merits for others who might be interested in your data.

Authors' response: Thank you for your comments and suggestion. We have made a supplementary table (Table S1 and S2) showing the number of raw sequence, OTUs, Chao I richness, and Shannon index in the Supplementary File.

The high-throughput sequencing data were submitted to GenBank on the basis of individual sample. Each sample has an unique accession number, e.g., the accession number of T90 is SRR1295592.

(2) Table 1 is a great supplement with very useful background information. However, this table needs to be re-organized and it is certainly not well adjusted for publication.

Authors' response: Thank you for your suggestion. We have split Table 1 in to Table 1 (Reservoir characteristics of Lu and Liu field block) and Table 2 (Chemical properties of the water samples obtained from Lu and Liu field block).

(3) Please clearly define the two different strategies, particularly ***the drawback for the indigenous microbial flooding technique***. The authors state that "Indigenous microbial flooding technique improves oil recovery by introducing oxygen and salts through water-based injection to stimulate reservoir microorganisms. ***Despite the validity in field trial, this technology also has some limitations, in particular, instability during microbial flooding process." It is not clear here***. The injection of exogenous microorganisms and of ex-situ products has drawback due to the sieve effect of strata on microbial cells. So the oxygen and salts are introduced to stimulate the indigenous microorganisms. However, to my present understanding the migration of reservoir microorganisms would be subjected to sieve effect of strata belowground. So ***what does the author really mean about "instability during microbial flooding process for the indigenous microbial flooding technique***".

Authors' response: We agree with your comment that it is needed to clearly define the drawback for the indigenous microbial flooding technique.

Petroleum reservoirs represent special environments underground with multiphase fluids of oil, gas, water, and high pressure. Because of the nutrients concentration, the sieve effect of reservoir strata, and the stochastic processes, microbial assemblages in reservoir strata may be different. Therefore, the "instability during microbial flooding process for the indigenous microbial flooding technique" refers to that microbial community diversification may have a significant influence on oil displacement efficiency.

To avoid misunderstanding, we made a correction in the revised manuscript: 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).

(4) *The lack of control. Please clearly state that it is difficult to document the indigenous microbial community structure" somewhere in the text.*

Authors' response: Thank you for your suggestion. We have discussed the issue on the indigenous microbial community in the discussion section.

The revised section is: 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.

(5) I personally feel that ***backflow water sample cannot be used as a control*** (see comments below). In addition, ***water-injection with some nutrients is said to stimulate the growth of certain indigenous microbial functional guilds***. But it is rarely discussed.

Authors' response: We agree with you that the backflow water sample cannot be used as a control. In the manuscript, the backflow water samples were actually used to state the similarities and differences in microbial communities in injected water and those in water samples retrieved form the reservoir strata.

According to your suggestion, we have discussed the relationship of nutrients injection, stimulation of certain indigenous microbial populations, and the oil displacement in the manuscript. The revised section is: 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).

(6) At the place of first appearance, please define what is “injection and production water samples of water-flooding petroleum reservoirs”. Most people would not have time reading every words of your manuscript. Therefore, ***it would be very useful if you could clearly mention what the purpose is for injection well and production well*** in line 23-24(very briefly).

Authors’ response: Thank you for your suggestion. We have described the purpose for community analysis of injection well and production wells in the 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 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.

(7) Line 31. This statement seems to be contradictory to the previous one. If the community structure exhibited large differences between the injected and produced water samples (Line 31), a large number of microbial populations would not have been shared in the conglomerate reservoir (in contrast to what have been observed in Line 29-30).

Authors’ response: Thank you for your suggestion. We have made a correction in the revised manuscript: 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. 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 injected and

produced water samples.

(8) Line 37. The term “relative abundance” could be removed

Authors’ response: We have made a correction in the revised manuscript.

(9) Line 38. Because the authors do not determine oxygen concentration, it cannot be placed.

Authors’ response: Thank you for your suggestion. We have made a correction in the revised manuscript.

(10) Line 137. Please clearly state what “Backflow” is. Is this water flowing upward through an independent well?

Authors’ response: Backflow refers to injected water flowed upward through injection well. To avoid misunderstanding, we made a correction in the revised manuscript: 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.

(11) Line 200. This section needs to be described and discussed in greater detail.

Authors’ response: According to your suggestion, we have described and discussed the quantification results of community abundance in detail.

(1) The results of qPCR indicated that the copy number of bacterial 16S rRNA in the injected water was 8.25×10^{6} copies ml^{-1}, while 1.5×10^{6} to 2.75×10^{6} copies ml^{-1} in the produced water samples. 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 injected water and 8.5×10^{3} to 5.75×10^{4} copies ml^{-1} in the produced water samples.

(2) 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.

(3) 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.

(12) Line 368. I think ***there is no solid evidence in support of the sequencing depth effect in this study***. The authors need to be careful about this statement. Provided that the number of sequence reads is more than 3000, the core structure of microbial community might be captured regardless of sequencing methods.

Authors' response: In this section, we intended to interpret that deeper sequencing is needed for detecting rare archaeal populations when using Miseq-sequencing based on bacterial and archaeal universal primer set 515f and 806r.

To avoid misunderstanding, we have made a correction in the revised manuscript: 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 results imply that deeper sequencing is needed for detecting rare archaeal populations using Miseq-sequencing based on

universal primer set 515f and 806r.

(13) Line 467. In Conclusion section, please clearly state what is the main difference, i.e., the taxonomic identity of microorganisms that exhibited large differences. In addition, the term of “structure” might be used throughout the manuscript because the authors conducted additional experiments to quantify the population size.

Line 475. The authors need to be careful about this conclusion, because the water backflow would most likely be different when compared to injection water. The reason is that the soil and subsurface sediment might serve as filters, leading to the difference between injected water and backflow water? Please carefully explain this point.

Authors’ response: Thanks for your suggestion. We have revised the conclusion section according to your suggestion.

The revised conclusion is: Using high-throughput sequencing, this study revealed the similarities and differences in microbial communities in the injected water and 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 the injected water. However, microbial community compositions exhibited large differences between the injected produced water samples. 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 injected water, while microaerophilic bacteria, facultative anaerobes, and anaerobes dominated the reservoir strata.

(14) All figure legends need to be re-organized. It is expected that your figure legends will be quite detailed and very precise. In fact, from the figure title and the axis labels of a graph/table the reader should be able to determine the question being asked, get a good idea of how the study was done, and be able to interpret the figure without

reference to the text.

Authors' response: Thanks for your suggestion. We have revised the figure legends according to your suggestion.

Authors' responses to the reviewers' comments and suggestions

Anonymous Referee 1

Received and published: **3 Jan. 2015**

General comments:

The manuscript entitled "The shift of microbial population composition accompanying the injected water flowing in the water-flooding petroleum reservoirs", by P.K. Gao and colleagues, describes the fluctuation of microbial communities along with injected water flowing into reservoir strata and production wells.

Although there are some interesting findings in this manuscript, it is not acceptable in its present form. I think that experimental design needs to be further improved, some of the interpretations should be better qualified and perhaps even modified. Occasionally, the text may be misleading, since the usage of English is sometimes not adequate. Hence, the manuscript should be carefully revised.

Our response: We thank you for the time and thought you put into reading our manuscript, and for your helpful suggestions for improvement.

Based on the comments and suggestions, we have made a readjustment to improve the precision of our conclusions based on the data obtained in this study.

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.

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, detailed studies about the effects of the microbial communities in injected water on those in reservoir strata remains poorly understood. 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? Therefore, this study investigated the similarities and differences in microbial communities in the injected water and reservoir strata in a homogeneous sandstone reservoir and a heterogeneous conglomerate reservoir.

We found the bacterial and archaeal communities in the reservoir strata have high concentrations, which are similar with those in the injected water. However, microbial community compositions exhibited large differences between the injected produced water samples. 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.

Major comments:

Question 1*: * English should be significantly improved. This manuscript suffers from grammar errors and poor writing, particularly, in the Results and Discussion section.*

Our response: We have carefully revised the manuscript according to the comments. To improve the quality of the paper, the revised manuscript has been edited by an English Language Editing Service.

Question 2*: * The Title should be reconsidered so that it can directly present the novel findings to the readers.*

Our response: Thanks for your suggestion. The title of the manuscript has been revised to “Differences in microbial community composition between injection and production water samples of water-flooding petroleum reservoirs”.

Question 3: ** Experimental design: the authors chose a sandstone reservoir and a conglomerate reservoir in a Chinese typical oilfield, and analyzed the microbial population composition in injected water and produced water samples by using high-throughput sequencing technology, in order to test whether microbial populations in injected water could pass through oil-bearing strata. Overall the*

approach is straightforward. However, there are three major shortcomings in this study:

(i)The lack of control: These two kind of old well groups have water flooded for 13- and 30-years. The negative controls are missing. The indigenous microbial community in the same oil-bearing strata cannot be overlooked. A better way is to determine oilfield water samples from newly drilled well for comparison in the same oil-bearing block.

Our response: We agree with your comment that negative controls, which may be water samples from newly drilled wells without water-flooding in the same oil-bearing block, are necessary to provide background information on indigenous microbial populations. Unfortunately, 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, it is difficult to obtain the reliable information on the indigenous microbial community. We have discussed this issue in discussion section in the revised manuscript.

We realize that it is less rigorous to delineate the transport of microbial populations in reservoir strata by only detecting the shared microbial populations in both injection and production wells using the 16S rRNA sequencing method, because the method is not able to distinguish whether the species detected in the produced water are the same as those in the injected water. To improve the precision of the conclusions based on the data obtained in this study, we have made a readjustment, which emphasizes the similarities and differences in microbial communities in the injected water and reservoir strata. We hope this revision will meet with your approval.

(ii) Some important geological parameters are missing: In this manuscript, by comparison of Lu and Liu field block reservoirs, the authors concluded that injected water can pass through reservoir strata, but the reservoir heterogeneity, sieve effect of strata and dissolved oxygen affect the microbial migration. However, geological parameters such as source rocks and oil sources of Lu and Liu field, the characteristics of crude oil (heavy oil or light oil), reservoir pressures, the depths of Lu and Liu oil-bearing strata are missing. If there are great differences between Lu and Liu field block in these parameters, the comparison does not make sense.

Our response: Thanks for your suggestion. We have added the geological parameters to the manuscript. The reservoir characteristics are listed in Table 1 and Table 2.

Although there are some differences in the reservoir characteristics of the Lu and Liu field blocks, the two reservoirs are both located in the Junggar Basin of the Xinjiang Uygur Autonomous Region, Northwest China. 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 has a higher content of saturates and aromatics, which facilitate the growth of hydrocarbon-degrading bacteria. The salinity of Lu block is approximately 11, 000 mg/L, which is similar to the value at Liu block. The cations and anions in the water samples in each block are similar, with a lower total nitrogen and total 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.

(iii) Environmental parameters such as dissolved oxygen (DO) or oxidation-reduction potential (ORP), subsurface temperature, and the composition of organic matter of crude oil (as substrate for fermentative bacteria and methanogenic archaea) have not determined in this study. These parameters are crucial for the fluctuation of microbial communities besides injected water. Therefore, it is also necessary to include these parameters in PCoA analysis. Without these information, the conclusion was untenable.

Our response: According to your suggestion, we have listed the subsurface temperature of the two reservoirs, the component of crude oil, and the concentrations of cations and anions in Table 1.

The subsurface temperatures of the two reservoirs are 37^{o}C and 22.6^{o}C, respectively. The concentration of nutrient characteristics was also measured, including crude oil properties, total nitrogen, total phosphorus, and ion concentration of formation water (Table 1). The ratio of saturates in the two reservoirs are 71.9% and 61.94 %, respectively, while the aromatic content is 14.85% and 11.24%. The resin and asphaltenes content is low. Among them, saturates and aromatics can be used as a carbon source for hydrocarbon-degrading bacteria (HDB), and some anaerobes, such

as sulfate-reducing bacteria. The salinity of Lu block was approximately 11, 000 mg/L, which was similar to the value of the Liu block. The cations and anions among the water samples in each block were similar, with a lower total nitrogen and total phosphorus content, which are essential for the survival and growth of microorganisms. The lack of nitrogen and phosphorus implies a low metabolism level of microorganism. We have provided this information in section “2.1 Sampling locations”.

Unfortunately, in situ oxygen concentrations were not measured at this time. However, microbial populations were clustered to highlight the populations that showed the most variability between the injected and produced water samples. We found that aerobic bacteria, including *Marinobacterium*, *Paracoccus*, *Ochrobactrum*, *Sphingomonas*, *Alcanivorax*, and *Azospirillum*, were detected in higher quantities in the injected water, while *Pseudomonas*, *Rhizobium*, *Arcobacter*, *Halomonas*, *Spirochaeta*, *Bacillus*, *Thauera*, *Halomonas*, and *Bacteroides* are microaerophilic bacteria, facultative anaerobes or anaerobes, that were dominant downhole of the injection and production wells. We think these data reflect the influence of dissolved-oxygen on microbial community diversification. Furthermore, Unifrac PCoA analysis was performed based on microbial OTUs abundance and phylogenetic relationships to extract and visualize the few highly informative components of variation from complex, multidimensional data. The results suggest the relative similarity and diversification of microbial communities in the injection and production wells. In the biplot, samples with similar community were placed at a close proximity, while samples with different community were placed at a comparatively decentralized position.

Question 4: ** Materials and methods: Amplicon sequencing was performed on two kinds of highthroughput sequencing platform (GS FLX+ and Miseq). Apparently, the outcomes must be different. Did the authors want to make a comparison? I cannot find any clue in this manuscript.*

Our response: Thank you for your suggestion. We have made a relevant discussion in the manuscript. The revised section is as follow: 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.

Question 5: ** Discussion: There is no "going home" feeling in this part. Too many hypothesizes were demonstrated. In situ DO in injected and production water and the composition of crude oil should be determined firstly. Then, the content of the true part of the "Discussion" should be carefully revised accompanying with more related new references.*

Our response: The concentrations of nutrient factors including crude oil properties, nitrogen, phosphorus, and ion concentration of formation water, were added in the manuscript (Table 1 and Table 2).

The "Discussion" has been carefully revised accompanying with more related new references. Please refer to the revised manuscript for the detailed correction.

Question 6: ** The conclusion is too long and should be carefully rewritten.*

Our response: According to your suggestion, we have carefully rewritten this section. The revised conclusion section is as below: Using high-throughput sequencing, this study revealed the similarities and differences in microbial communities in the injected water and 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 the injected water. However, microbial community compositions exhibited large differences between the injected produced water samples. 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 injected water, while microaerophilic bacteria, facultative anaerobes, and anaerobes dominated the reservoir strata.

Minor comments:

Question 1: ** Page 16774 Line 5-8: "The results indicated that the injected water: : : in this reservoir." This sentence needs to be rephrased.*

Our response: Thanks for your comment. We have made a readjustment to improve the preciseness of our conclusion based on data obtained in this study. The abstract was revised accordingly. The revised abstract is as follow: 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 injected and produced water samples 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

injected and produced water samples. 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.

Question 2: ** Page 16777 Line 10-12: "All the injected and produced water samples were all collected from the wellhead of injection and production wells by the field personnel of PetroChina." Once the samples were collected, how long would be the genomic DNA extracted?*

Our response: The transportation from Xinjiang Oil Field to Tianjin is seven days. To avoid misunderstanding, we have made correction in the manuscript. The revised section is as below: "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 the laboratory, 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)."

Question 3: ** Page 16779 Line 13-14: why do not quantitate archaeal populations? It would be helpful for the interpretation of methanogenic community later.*

Our response: Thank you for your suggestion. According to your suggestion, we have quantitated the number of archaeal populations using genome DNA preserved.

Question 4: *Page 16780 Line 3-4; Page 16781 Line 2-3: The diversity of the microorganisms in the sandstone reservoir (249-538) is much lower than it in the conglomerate reservoir (51273-128980). Why? In this respect, the geological and environmental condition of Lu and Liu field might be very different from each other.*

Our response: Thank you for your comment. We have made correction in the manuscript. The truth is that 4016–5060 bacterial sequences and 2688–2857 archaeal

sequences were obtained by pyrosequencing in the sandstone reservoir, while 52719 to 129106 16S rRNA gene sequences were obtained by Miseq-sequencing in the conglomerate reservoir. We have made discussion in the manuscript.

Question 5: ** Page 16781 Line 17: The word of "botained" should be revised to "obtained".*

Our response: Thanks for reminding us. We have made a correction in the manuscript.

Question 6: ** Page 16784 Line 17: "LR, 2010" should be revised to "Brown, 2010".*

Our response: Thanks for reminding us. We have made correction in the manuscript.

Question 7: ** Page 16791 Table 1: I guess the temperature was the surface temperature. The temperature of oil-bearing strata should also be given.*

Our response: Thank you for your comment. The listed temperature (37 ^{o}C and 22.6 ^{o}C) is the subsurface temperature of the two reservoirs.

Question 8: ** Page 16793 Figure 2: The color representing a designated microbial class should be consistent so that it is easy for comparison. (e.g. The color of Methanococci is blue in Fig. 2aII but red in Fig. 2bII.)*

Our response: Thank you for your suggestion. We have made correction in the revised manuscript.

Comments from reviewer Prof. Jidong Gu

Received and published: **2 Jan. 2015**

General comments:

This published paper on 'The shift of microbial population composition accompanying the injection water flowing in the water-flooding petroleum reservoirs' by Gao et al. certainly shows some descriptive information on the possible transport of microorganisms through oil reservoir subsurface sandstone materials.

Our response: Thank you for reviewing our manuscript, and your constructive comments and suggestions. Based on the comments and suggestions, we have been revising our manuscript in an effort to improve it.

***Question 1**: * I have to say that the title does not fit with the data obtained because the samples did not include a nonintervention control to allow assessment of the indigenous population for a meaningful comparison. Without this critical sample and information, the transport of bacteria is a claim not supported by convincing data*.

Our response: We thank you for the constructive advice, which has been important and valuable in improving the manuscript quality.

We agree with your comment that it is less rigorous to delineate the transport of microbial populations in reservoir strata by detecting the shared microbial populations in both injection and production wells using a 16S rRNA sequencing method. Unfortunately, because the two reservoirs have been long-term water-flooded, the indigenous microbial community in the subsurface might have been disturbed. Thus, it is difficult to obtain the reliable information on the indigenous microbial community. We have made a readjustment to improve the precision of our conclusions based on the data obtained in this study.

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.

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, detailed studies about the effects of the microbial communities in injected water on those in reservoir strata remains poorly understood. 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? Therefore, this study investigated the similarities and differences in microbial communities in the injected water and reservoir strata in a homogeneous sandstone reservoir and a heterogeneous conglomerate reservoir.

We found the bacterial and archaeal communities in the reservoir strata have high concentrations, which are similar with those in the injected water. However, microbial community compositions exhibited large differences between the injected produced water samples. 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.

As a result, we revised the manuscript title to “Differences in microbial community composition between injection and production water samples of water-flooding petroleum reservoirs”.

Question 2*: * In the text, the differences of detected pyrosequences between injection water and production water were used as the key variables to delineate the transport (migration) of microorganisms, a major shortcoming with this approach is that some microorganisms will not survive the subsurface environmental conditions due to lack of oxygen, nutrients etc. the approach used in this research plan should be reconsidered.*

Our response: Thanks for your suggestion. Since Bastin et al. firstly isolated

sulfate-reducing bacteria (SRB) from reservoir in 1926 (Bastin, 1926), culture-independent methodologies, such as fluorescence in situ hybridization (FISH), denaturing gradient gel electrophoresis (DGGE), terminal restriction fragment length polymorphism (T-RFLP), quantitative PCR, 16S rRNA clone libraries, and high-throughput sequencing have revealed diverse microbial populations inhabiting petroleum reservoirs. Although these methods could not delineate survive and activity of reservoir microbial populations, culture-dependent methods combination with culture-independent methods have demonstrated the existence and activity of hydrocarbon-degrading bacteria (HDB), nitrate-reducing bacteria (NRB), sulfate-reducing bacteria (SRB), and methanogens in injection and production water samples. On the other hand, even there is no metabolic activity in the subsurface environmental conditions due to lack of oxygen and nutrients, these populations may be dormant in reservoir.

To improve the precision of our conclusions based on the data obtained in this study, we have made a readjustment. The study was performed to illustrate the relationship shared by microbial communities in the injection and production water samples. We think that the high-throughput sequencing can distinguish the differences between microbial communities in water samples collected from wellhead or downhole of injection wells, and production wells.

Question 3*: * First of all, I am sorry to say that the quality of this manuscript writing is low and it is hard to read the text for accurate meaning and the precise information. The writing needs extensive efforts and time to revise to reach to a reasonable level of acceptance. Authors must work hard on this and serious because the results of the information can be compromised seriously when the statements cannot be comprehended well enough by reading*.

Our response: Thanks for your suggestion. We will work hard on writing in this and our future paper. To improve the quality of the paper, the revised manuscript has been edited by an English Language Editing Service.

Question 4*: *There is little or any disagreement now that oil reservoirs have*

indigenous population of microorganisms, but non-indigenous microorganisms are introduced into the reservoir systems when water flooding is introduced. It is always a big challenge to obtain the truly indigenous population of microorganisms in the reservoirs because of the difficulties involved in non-contamination sampling of the subsurface environment without any potential contamination. In a similar but different aspect, the physical characteristics of the subsurface materials, either heterogenous or homogenous as stated in this paper is also a term of personal choice here than substance because of their natural origin and heterogeneity no matter called heterogenous or homogenous. Heterogeneity is the true nature of such materials. Therefore, I have concern on the choice of 'homogeneity' and 'hererogeneity' simply based on the average permeability values because this value is an average numerical number, which cannot be used reliably for transportability of bacteria. Considering the differences in permeability between the two blocks, there should be no disagreement on bacteria can be transported in both subsurface systems, but the rate of transport may be different. If this is the case, what is the key scientific information that can be extracted from the selection of the 2 blocks in this investigation? If the injection of water had only started with this study, the collected water/oil samples can be of some meaning interpretation, but I do not think such is the case with this set of production wells.

Our response: Thanks for your comment. As you pointed out, the two reservoirs have been water flooded 13- and 30-years, non-indigenous microorganisms might be introduced into the reservoir systems, and the indigenous microbial community in the subsurface might have been disturbed. Thus, it is difficult to obtain the reliable information on indigenous microbial community, even if we obtained water samples from newly drilled well.

We also realized that it was less rigorous to delineate the transport of microbial populations in reservoir strata by detecting the shared microbial populations in both injection wells and production wells using a 16S rRNA sequencing method, because it is not able to demonstrate whether the species detected in produced water are the same ones in the injected water.

To avoid misunderstanding, we made a discussion in the manuscript: <u>This study investigated the relationship shared by microbial communities in injected water and</u>

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 water samples, injecting labelled strains containing marked gene (e.g., green fluorescent protein coded gene) into reservoirs may bring novel insight and greater predictive power.

Question 5*: *The 'approximately 30-45 days and 7-10 days, respectively' – I have no way of knowing what do they refer to by the sentence because no designation was offered. I may assume they are associated with the heterogenous and homogenous reservoirs, but such assumption should not be the responsibility of the readers and they must be clearly stated by the authors to avoid any misunderstanding.*

Our response: Thanks for your suggestion. We have revised the sentence as "Tracer technique indicated that the time intervals for injected water from injection well flowing into neighboring production wells are approximately 30–45 days in the sandstone reservoir, while 7–10 days in the conglomerate reservoir."

Question 6*: *Sampling procedures were inadequately described and I am especially troubled by the statement ': : :by the field personnel of PetroChina.' because the quality of the samples may be compromised for one. In addition, how can the authors interpret the results when they are not involved in the in situ sampling to know the*

detail steps involved and the effects on the results obtained?

Our response: Thanks for your suggestion. Actually, we involved in the in-situ sampling, and the sampling process was assisted by the field personnel of Oil Field.

To avoid Misunderstanding, we have made correction in the revised manuscript. The revised section is: “All the injected and produced water samples were collected randomly from sampling valves located on the wellhead. The water samples were completely filled into 15 L sterilized plastic bottles, which were immediately capped and sealed to avoid contamination and oxygen intrusion.”

***Question 7**: * Further on the sampling for concentration of bacterial cells, oil/water mixture should separate the oil from the mixture and then concentrate the cells from water phase or both oil and water phases. This detailed information show the understanding of the system you are dealing with and the quality of the cells you would be obtained. Are there any differences in terms of the composition and richness of microbial groups associated with the oil and water phases? Why was the oil phase not treated for extraction of DNA in the similar way as water phase? Actually, recent publication(s) has/have some information on this topic and you should also cited the work here. Were there any quality controls in the extraction of genomic DNA and PCR amplification?*

Our response: This is a good inspiration to our future research in this direction. According to your suggestion, we have looked through the recent publications, which compared the similarities and differences of microbial communities in oil phase and water phase.

We have made a relevant discussion in the manuscript: 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.

***Question 8**: * 'In the sandstone reservoir' – I do not agree with you to have such a statement and claim simply because there is no strictly control, which did not have any water flooding to show the indigenous population and composition. If the objectives of this study are on migration of microorganisms in subsurface sandstone, I do not think the experimental design can answer the questions effectively. This is a key point in Discussion, I have strong reservation in accepting this. The high-throughput used can be sensitive for detection of microorganisms in samples, but they do not answer the transportability of microorganisms without careful planning, selection of samples (including subsurface) and the analysis involved.*

Our response: We agree with your comment that it is less rigorous to delineate the transport of microbial populations in reservoir strata by detecting the shared microbial populations in both injection wells and production wells using a 16S rRNA sequencing method.

To improve the preciseness of the manuscript, we have revised the manuscript title as "Differences of microbial community composition between injection and production water samples of water-flooding petroleum reservoirs". 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. 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 injected and produced water samples. 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.

***Question 9**: * How can you link the microbial groups detected and the possible physiological function in the oil reservoirs? What are the sources of Bcteroides in the production water samples? From the information of archaea detected, which kind of methanogenic metabolism is responsible for CH_4 production?*

Our response: A number of fermentative microorganisms have been isolated from high- and low-temperature oil reservoirs. Many microorganisms in this group possess dual fermentative and respiratory metabolic abilities and could theoretically use both strategies for their in situ growth and survival (Youssef et al. 2009). *Bacteroidetes* include diverse mesophilic fermentative microorganisms. Grabowski et al. first isolated *Petrimonas* within the phylum *Bacteroidetes* in 2005 as the first member of this phylum to be isolated from oil reservoirs (Grabowski et al., 2005). Recently, studies based on 16S rRNA gene-based analysis have also revealed the existence of putatively fermentative members of the genus *Bacteroides* (Grabowski et al., 2005b; Youssef et al. 2009). These fermentative microorganisms may play an important role in reservoir ecosystems, in particular, providing substrates for methanogens to produce methane.

According to your suggestion, we have classified the obtained archaeal taxa based on the reported methyltrophic, acetoclastic, and CO_2-reducing methanogens (Liu, 2008). As reported in previous research (Zhao et al., 2012;Wang et al., 2012), the archaea identified in both reservoirs were overwhelmingly methanogens including methyltrophic, acetoclastic, and CO_2-reducing methanogens. Among them, methyltrophic and CO_2-reducing methanogens dominated both reservoirs.

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%). Among them, Methanosaeta uses only acetate to produce CH_4. *Methanomicrobia* made up 84.03% of 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.

Similarly, *Methanobacteria*, *Methanococci*, and *Methanomicrobia* composed 64.3%–94.6% of the archaeal communities in the conglomerate reservoir. Compared with the injected water collected from the wellhead of the injection wells (T86-0 and T93-0), more *Methanomicrobia* was detected downhole of the 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 use H_2 and formate as carbon sources to produce CH_4.

Question 10*: * The Conclusions is too lengthy and shortening is necessary to show the most significant information of this research if any. References should be updated more extensively to include the current published papers to enrich the information reported here.*

Our response: Thanks for your suggestion. We have carefully rewritten the conclusion: Using high-throughput sequencing, this study revealed the similarities and differences in microbial communities in the injected water and 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 the injected water. However, microbial community compositions exhibited large differences between the injected produced water samples. 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 injected water, while microaerophilic bacteria, facultative anaerobes, and anaerobes dominated the reservoir strata.

The manuscript has been carefully revised referring to more current published papers. To improve the paper quality, the revised manuscript has been edited by English Language Editing Service before resubmission.

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

批注 [A1]: According to the reviewer' suggestion, the title was reconstructed to precisely reflect the main content of this article.

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 injected and produced water samples 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 injected and produced water samples. 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

批注 [A2]: To be more precise, we have rewrote the abstract.
In this section, we introduce the research background, and described the purpose for community analysis of injection well and production wells.

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

批注 [A3]: We provided information about the potential effects of injected water on microbial community in reservoir strata.

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.

批注 [A4]: We compared our study with more recent publications about microbial communities in injected water and produce water samples.

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, interwell 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).

批注 [A5]: According to the reviewer' suggestion, more physical and chemical properties were added.

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 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 (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.

批注 [A6]: Backflow is described in detail.

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.

批注 [A7]: According to the reviewer' suggestion, the section was described in detail.

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).

批注 [A8]: The high-throughput sequencing data were submitted to GenBank on the basis of individual sample. Each sample has an unique accession number, e.g., the accession number of T90 is SRR1295592.

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. 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 the injected water was 8.25×10^6 copies ml^{-1}, while 1.5×10^6 to 2.75×10^6 copies ml^{-1} in the produced water samples. 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 injected water and 8.5×10^3 to 5.75×10^4 copies ml^{-1} in the produced water samples.

批注 [A9]: We have described and discussed the quantification results of community abundance in detail.

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.

批注 [A10]: According to the reviewer' suggestion, we classified the obtained archaeal taxa based on the reported methyltrophic, acetoclastic, and CO_2-reducing methanogens

批注 [A11]: According to the reviewer' suggestion, we classified the obtained archaeal taxa based on the reported methyltrophic, acetoclastic, and CO_2-reducing 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. 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}.

批注 [A12]: We have described and discussed the quantification results of community abundance in detail.

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

290 H_2 and formate as carbon sources to produce CH_4.

291

292 **3.3 Shared microbial populations between injected water and reservoir strata**

293 The shared microbial OTUs and genera between communities in the injected and
294 produced water samples were investigated using Venn diagrams, histograms, and
295 heatmap. Based on the Venn diagrams, 16.3%–32.81% of bacterial OTUs and
296 13.73%–51.61% of archaeal OTUs were shared between the injected water and each
297 produced water sample in the sandstone reservoir (Fig. 3a). These shared bacterial
298 OTUs accounted for 4.6–24.71% of the total bacterial community in the injected
299 water, and 43.23–76.18% in each produced water sample (Fig. 4a I). Furthermore,
300 only 13 bacterial and 3 archaeal OTUs were shared by both the injected and produced
301 water samples (Fig. 3a). Accordingly, the shared genera only accounted for 2.26% of
302 the bacterial community in the injected water, but dominated each production well
303 (12.02%–36.5%; Fig. 4a II). Similar to the bacteria, three archaeal genera belonging
304 to *Methanobacterium*, *Methanococcus*, and *Methanolobus* were detected in the
305 injected and produced water samples, comprising 13.58% of the total archaea in the
306 injected water, and 90.4%–96.89% in each of the produced waters (Fig. 4a III).

307

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

314

批注 **[A13]:** According to the reviewer' suggestion, we classified the obtained archaeal taxa based on the reported methyltrophic, acetoclastic, and CO_2-reducing methanogens

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 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*, *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 injected water was distinct from that of the produced water (Fig. S3). Weighted PCoA distinguished the bacterial community of the injected water from that of the production wells, while communities of the production wells were placed at a comparatively decentralized position (Fig. 6a I). 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 II). 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).

批注 [A14]: To be more precision, we rewrote this section.

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 injected water and each production well 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 injected water and each produced water sample. 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

批注 [A15]: We discussed the the relationship of Miseq-sequencing depth and the quantification results of community abundance.

批注 [A16]: To be more precision, we rewrote this section.

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 by the injected water and produced water samples 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

批注 [A17]: We discussed the relationship of community abundance in the injected water and produced water samples.

批注 [A18]: We discussed the disadvantage 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.

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 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 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 downhole of injection and production wells.

批注 [A19]: To be more precision, we rewrote this section.

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

462 growth in addition to the water phase within the petroleum reservoir (Kryachko et al.,
463 2012;Kobayashi et al., 2012). Recent research has also compared microbial
464 communities in aqueous and oil phases of water-flooded petroleum reservoirs, and
465 found that the oil phase also harbored a large number of microorganisms, with large
466 differences in the bacterial community between the aqueous and oil phases of the
467 reservoir fluid (Wang et al., 2014). Therefore, simultaneous analysis of DNA
468 extracted from both aqueous and oil phases may provide a better understanding of the
469 microbial communities in injection and production water samples.
470
471 In summary, this study investigated the relationship shared by microbial communities
472 in injected water and reservoir strata in two long-term water-flooding reservoirs.
473 However, the results cannot provide any reliable information on the indigenous
474 microbial community. The indigenous microbial populations may be those in newly
475 drilled wells without water-flooding in the same oil-bearing block. However, the two
476 reservoirs have been water-flooded for decades. Due to the introduction of exogenous
477 microorganisms in injected water and other sources of contaminations by enhanced
478 oil recovery processes, determining whether a microorganism is indigenous to a
479 petroleum reservoir become increasingly difficult. This study implies that the number
480 of shared populations reflects the influence of microbial communities in injected
481 water on those in reservoir strata to some extent, and show strong association with the
482 unique variation of reservoir environments. However, it cannot make a conclusion on
483 the transport of microbial populations in the reservoir strata by detecting the shared
484 microbial populations in injected water and produced water samples using 16S rRNA
485 sequencing. To further investigate the relationship shared by microbial communities
486 in injection and production water samples, injecting labelled strains containing

批注 [A20]: According to the reviewer's suggestion, we have looked through the recent publications, and discussed the similarities and differences of microbial communities in oil phase and water phase.

批注 [A21]: According to the reviewer's suggestion, we have discussed the issue on the indigenous microbial community.

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 the injected water and 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 the injected water. However, microbial community compositions exhibited large differences between the injected produced water samples. 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 injected water, while microaerophilic bacteria, facultative anaerobes, and anaerobes dominated the reservoir strata.

批注 [A22]: According to the reviewer's suggestion, we discussed the disadvantage 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.
Additionally, injecting labelled strains containing marked gene (e.g., green fluorescent protein coded gene) into reservoirs is suggested for urther research on microbial diversification and transferability as injected water flows into reservoir.

批注 [A23]: To be more precision, we rewrote this section.

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.

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., Duan, R. Y., Liu, J. F., Yang, S. Z., Gu, J. D., and Mu, B. Z.: Molecular analysis of the microbial community structures in water-flooding petroleum reservoirs with different temperatures, Biogeosciences, 9, 5177–5203, 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., Chai, L. J., Banat, I. M., Zhang, X. T., Shu, F. C., Wang, Z. L., Yu, L. J., and Hou, D. J.: Microbial diversity in long-term water-flooded oil reservoirs with different in situ temperatures in China, Sci Rep, 2, 760, 10.1038/srep00760, 2012a.

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, 2012b.

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. 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. 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.

批注 [A24]: According to your suggestion, we have revised the figure legends to provide more detail information on how the study was done.

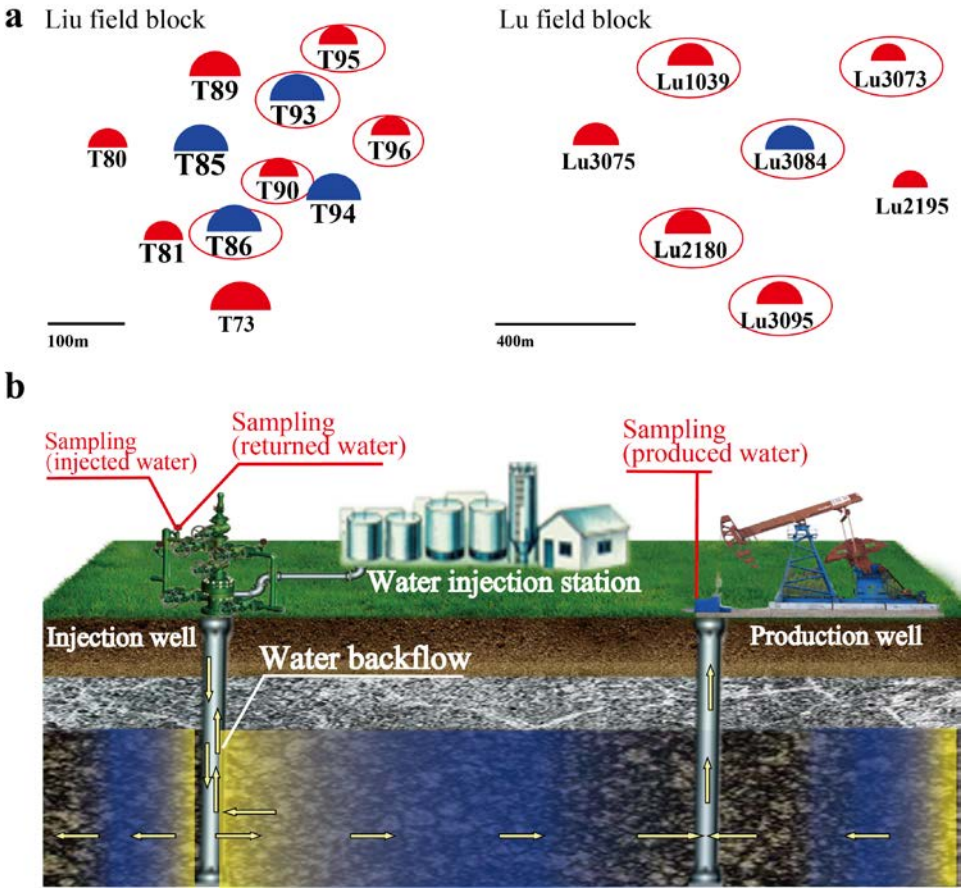
637 **Table 1** Reservoir characteristics of Lu and Liu field block

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

638

639 **Table 2** Chemical properties of the water samples obtained from Lu and Liu field
640 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



641

642 **Fig. 1**

643

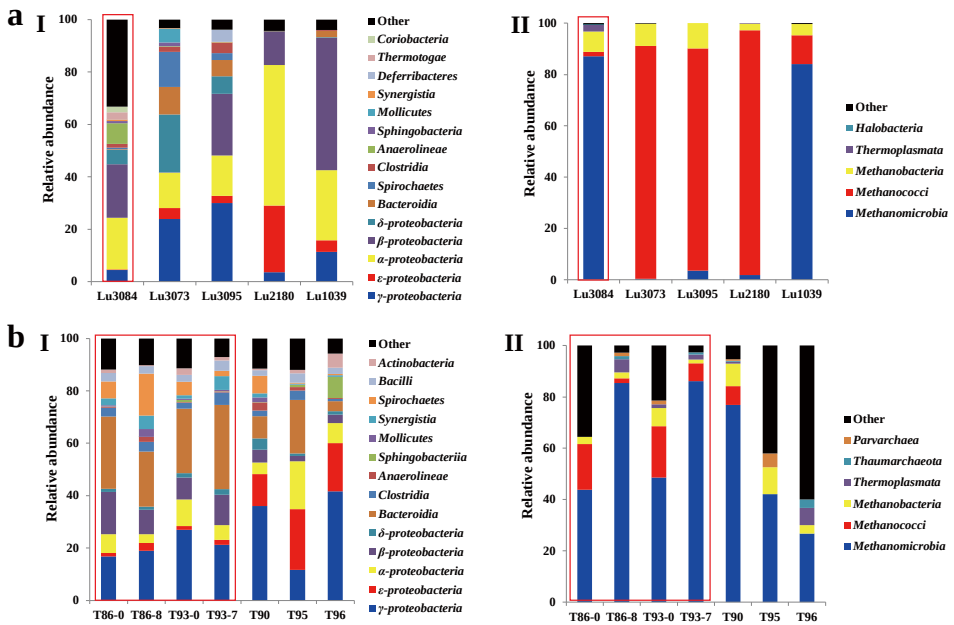

a
I
II
Relative abundance
100
80
60
40
20
0
Lu3084
Lu3073
Lu3095
Lu2180
Lu1039
Other
Coriobacteria
Thermotogae
Deferribacteres
Synergistia
Mollicutes
Sphingobacteria
Anaerolineae
Clostridia
Spirochaetes
Bacteroidia
δ-proteobacteria
β-proteobacteria
α-proteobacteria
ε-proteobacteria
γ-proteobacteria
Other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

Fig. 2

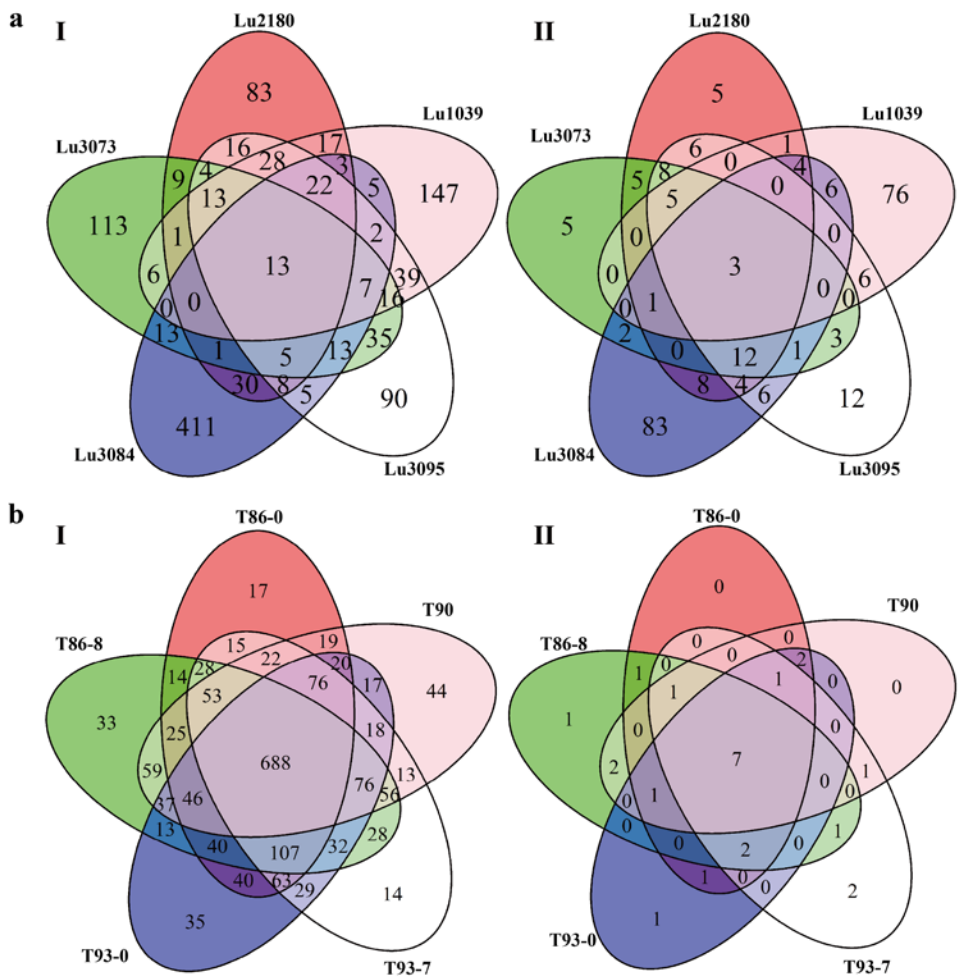
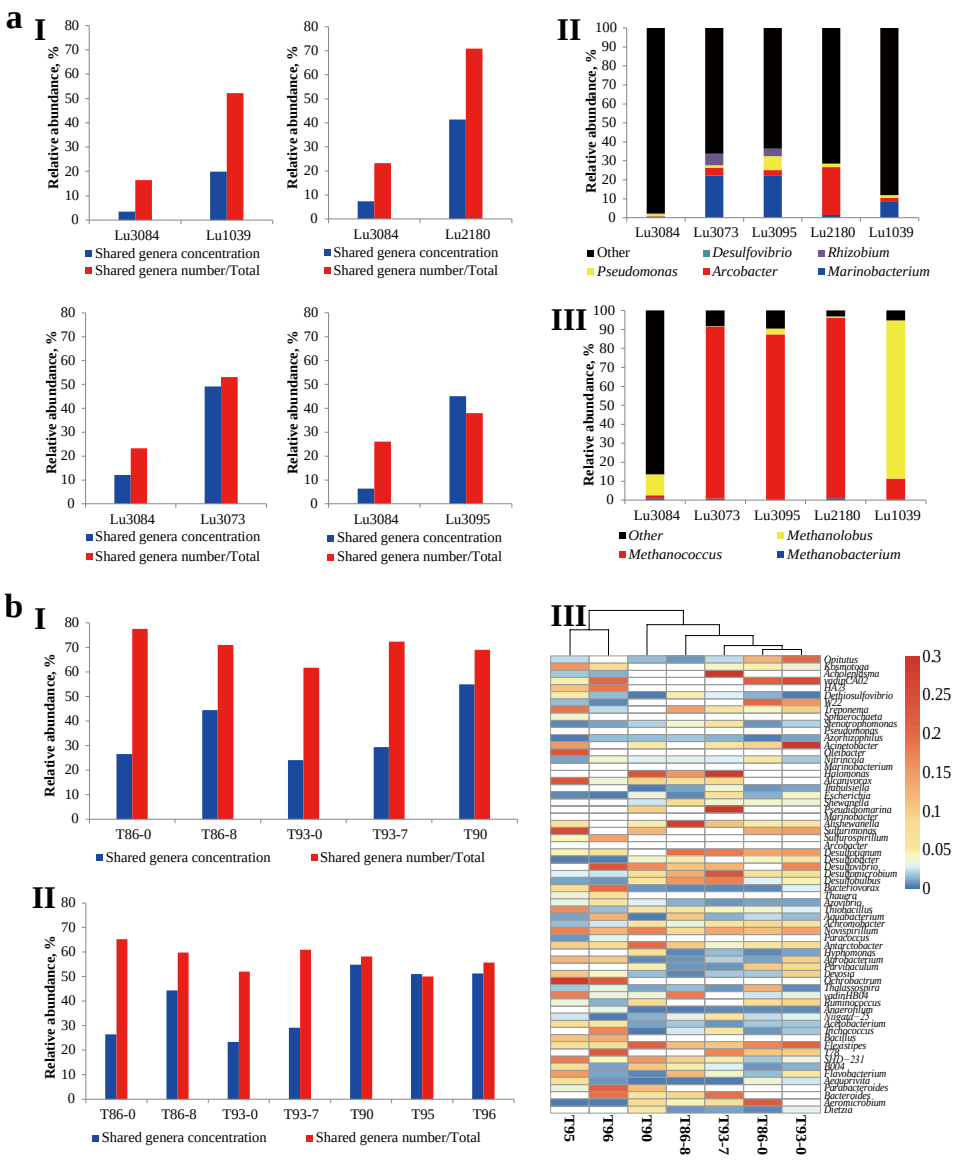

a
I
II
Lu2180
Lu1039
Lu3073
Lu3084
Lu3095
b
T86-0
T90
T86-8
T93-0
T93-7

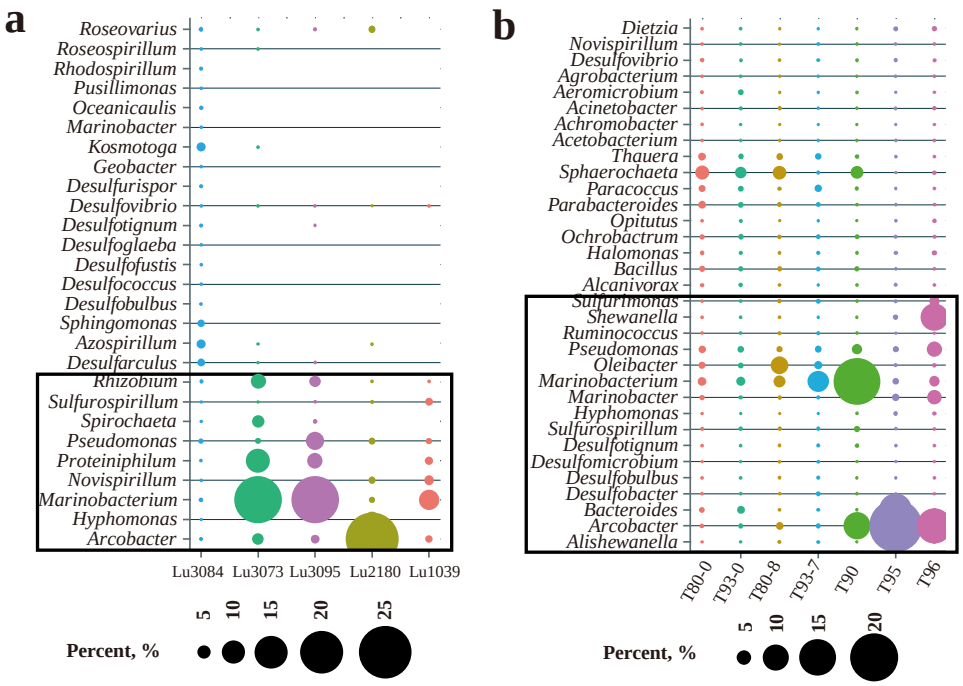
Fig. 3



650

651 **Fig. 4**

652



653

654 **Fig. 5**

655

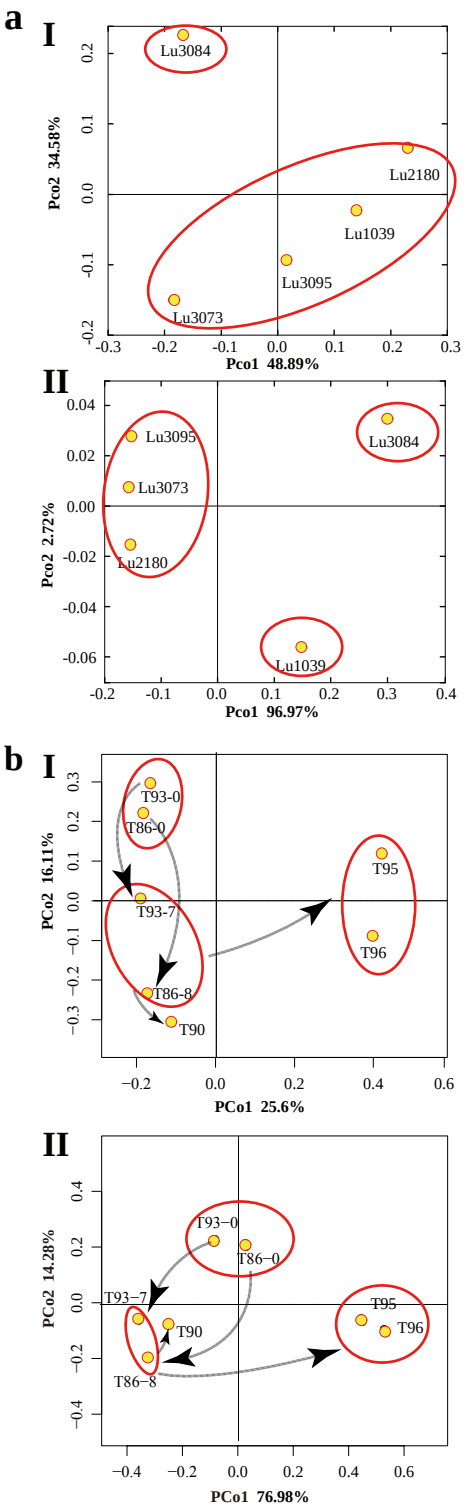

a
I
Lu3084
Lu2180
Lu1039
Lu3095
Lu3073
Pco2 34.58%
Pco1 48.89%
II
Lu3095
Lu3084
Lu3073
Lu2180
Lu1039
Pco2 2.72%
Pco1 96.97%
b
I
T93-0
T86-0
T95
T93-7
T96
T86-8
T90
PCo2 16.11%
PCo1 25.6%
II
T93−0
T86−0
T93−7
T95
T96
T90
T86−8
PCo2 14.28%
PCo1 76.98%

Fig. 6