



Distribution of Bacterial Communities in Petroleum-Contaminated Soils from the Dagang Oilfield, China

Xueke Feng¹ · Zhen Liu¹ · Xiaoqiang Jia^{1,2,3} · Wenyu Lu^{1,2,3}

Received: 11 November 2019 / Revised: 25 November 2019 / Accepted: 27 November 2019 / Published online: 19 December 2019
© The Author(s) 2019

Abstract

Diversity in bacterial communities was investigated along a petroleum hydrocarbon content gradient (0–0.4043 g/g) in surface (5–10 cm) and subsurface (35–40 cm) petroleum-contaminated soil samples from the Dagang Oilfield, China. Using 16S rRNA Illumina high-throughput sequencing technology and several statistical methods, the bacterial diversity of the soil was studied. Subsequently, the environmental parameters were measured to analyze its relationship with the community variation. Nonmetric multidimensional scaling and analysis of similarities indicated a significant difference in the structure of the bacterial community between the nonpetroleum-contaminated surface and subsurface soils, but no differences were observed in different depths of petroleum-contaminated soil. Meanwhile, many significant correlations were obtained between diversity in soil bacterial community and physicochemical properties. Total petroleum hydrocarbon, total organic carbon, and total nitrogen were the three important factors that had the greatest impacts on the bacterial community distribution in the long-term petroleum-contaminated soils. Our research has provided references for the bacterial community distribution along a petroleum gradient in both surface and subsurface petroleum-contaminated soils of oilfield areas.

Keywords Petroleum-contaminated soil · Dagang Oilfield · Bacterial community diversity · 16s rRNA Illumina sequencing · Environmental factor correlation analysis

Introduction

Microorganisms are the decomposers that play an indispensable role in the material exchange and energy operations of the ecosystem [1]. However, high energy efficiency, advanced petroleum extraction technology, and widely distributed oil wells promote the use of petroleum in industries.

Various petroleum industry activities have resulted in serious oil contamination in the environment [2]. Petroleum pollution affects the growth and reproduction of many kinds of microorganisms, resulting in changes in the soil microbial diversity [3].

In the soil, petroleum will bind soil particles together, which leads to low water and oxygen content [4]. Organic petroleum easily combines with inorganic nitrogen and phosphorus to limit the nitrification and dephosphorization of soil, slowing the circulation of nutrients. In addition, organic matter in petroleum can greatly increase soil organic carbon content. These changes in the physicochemical properties of the soil will change bacterial communities [5].

Previous studies on microorganisms were based on traditional separation culture [6–9], but most of the microorganisms in environmental samples cannot be cultured; therefore, it is difficult to objectively evaluate microbial diversity. Modern molecular biotechnology overcomes the shortcomings of pure culture methods and can estimate species diversity and richness from the genetic level, analyze species variation, and accurately understand the composition of microbial communities in the environment [10]. Based on

Xueke Feng and Zhen Liu have contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12209-019-00226-7>) contains supplementary material, which is available to authorized users.

✉ Wenyu Lu
wenyulu@tju.edu.cn

¹ School of Chemical Engineering and Technology, Tianjin University, Tianjin 300072, China

² Key Laboratory of System Bioengineering (Tianjin University), Ministry of Education, Tianjin 300072, China

³ Collaborative Innovation Center of Chemical Science and Engineering (Tianjin), SynBio Res Platform, Tianjin 300072, China

16S rRNA molecular marker technology, analysis of microbial diversity has become one of the important methods of microbial ecology research [11].

The soil alkane-degrading bacterial communities in the Karamay and Daqing Oilfields mainly consist of *Pedobacter*, *Mycobacterium*, and unknown alkB-harboring microorganisms [12]. The total petroleum hydrocarbon (TPH) concentration negatively correlated with soil microbial biomass carbon and positively correlated with microbial biomass nitrogen, but had no correlation with the evenness index of the Yellow River Delta region [13]. TPH was the most important factor that shifted the diversity of soil bacteria in Oman, USA [14]. Moreover, the relationship between the petroleum contamination gradient and bacterial diversity is still unclear [15].

Bioremediation has major advantages in utilizing the hydrocarbon-degrading microorganisms from environments to eliminate contaminants and has been considered as an efficient, environmental-friendly, and cost-effective technology [16]. It is essential and valuable to make in-depth investigations on microcommunities of the ecosystem. Cenozoic deposits and underground freshwater resources are abundant in the Dagang Oilfield of Tianjin, and they provide versatile environmental conditions and resources for soil microbes. We have focused on the diversity of the soil bacterial community and its correlation with environmental factors in the oil-contaminated soils of the Dagang Oilfield.

The purpose of this study is as follows: (1) to study the bacterial community structure in surface and oil-contaminated deep soils and compare it to nonpetroleum-contaminated soils; (2) to clarify the relationship between bacterial community diversity and the several environmental factors of petroleum-contaminated soil; and (3) to ascertain the microbial resources for further exploitations.

Materials and Methods

Geography of the Dagang Oilfield and Soil Collection

The Dagang Oilfield is located in southeastern Tianjin in between the Bohai Sea and the Hebei Province. It covers more than 20 districts, cities, and counties in Tianjin, Hebei, and Shandong Provinces. Since 1964, the oil production field area has covered approximately 18,716 km², including land, marine tidal flats, and shallow sea areas. The Dagang Oilfield contains abundant oil resources and very broad exploration prospects.

Soil samples were gathered from the Dagang Oilfield in January 2019. A total of 26 soil samples, including 3 pairs of nonpetroleum-contaminated surface and deep soil samples and 10 pairs of oil-contaminated surface and deep soil

samples, were collected. The sampling strategy was as follows: (1) the distance between each sampling point exceeded 100 m, and (2) different locations with various TPHs were sampled. Each site was marked as a square with a side length of 20 cm. Surface soils (5–10 cm) and deep soils (35–40 cm) were collected by a five-point sampling method. A total of 30 g of soil was collected from each site; afterward, a 2-mm filter was used to filter the soil sample. Each soil sample was divided into two parts: one for the determination of environmental factors and the other for DNA sequencing. The collected soil samples were air-dried and stored at –80 °C.

Physical and Chemical Analyses

The pH value was measured in a 1:5 (w/w) aqueous solution with a pH meter (PHS-3C, INESA, Shanghai, China). The electrical conductivity (EC) was measured in a 1:5 aqueous solution with a conductivity meter (FE30, Mettler-Toledo, Switzerland). The water content (WC) was detected by oven-drying fresh soil to a constant weight at 105 °C. The total organic carbon (TOC) content was measured by the dilution-heating method [17]. Total nitrogen (TN) content was determined using the Kjeldahl method [18]. Determination of total phosphorus and total potassium was done by the NaOH melting method [19].

DNA Extraction and Sequencing

The cetyltrimethyl ammonium bromide (CTAB) method was used to extract bacterial DNA from the soil (TIAN-amp Soil DNA Kit). Then, the sample was diluted to 1 ng/μL with sterile water. The qualitative analysis of DNA was conducted using agarose gel electrophoresis. The polymerase chain reaction (PCR) products were sequenced by an Illumina HiSeq. 2500 platform (250-bp paired-end reads).

Analysis of Illumina Sequencing Data

The raw sequence data were processed. Ambiguous data were removed, and the effective tags applied to the operational taxonomic unit (OTU) analysis were retained by QIIME (version 1.2.3.3). OTUs were chosen using the preferred open-reference OTU picking strategy in QIIME, and USEARCH was used to cluster sequences with similarity higher than 97% together to form an OTU. A representative sequence was chosen from each OTU, and the taxonomy was assigned to each of the representative sequences with a Ribosomal Database Project Classifier against Greengenes at a confidence threshold of 0.8 in QIIME. Representative sequences were aligned against the Greengenes core set using PyNAST. Then, using FastTree, a phylogenetic tree

was generated, which was used to calculate the weighted UniFrac distance.

Alpha diversity indices, which include observed species, the Shannon diversity index, and the Simpson diversity index, were calculated using QIIME. Principal coordinates analysis (PCoA) based on a weighted UniFrac distance matrix was calculated in the R Stats package (version 3.3.1). A heat map was drawn using the pheatmap package in R. Canonical correlation analysis (CCA), analysis of similarities (ANOSIM), and a mental test were conducted using the vegan package in R.

Data Availability

The original sequences of the research were deposited at the National Center of Biotechnology Information (NCBI) Sequence Read Archive under the accession number: PRJNA545042.

Results

Physical and Chemical Properties of Soils

The environmental parameters of 26 soil samples are shown in Table 1. The range of the TPH gradient was 0–0.4043 g/g, WC gradient was 0.1128–0.4070 g/g, pH gradient was 7.6–8.9, EC gradient was 735–8150 $\mu\text{s}/\text{cm}$, and TOC gradient was 6.880–172.000 mg/g. These environmental factors have a wide range of values, especially for TPH. Generally, the TPH of the surface soil in each pair was greater than that of the deep soil. Samples were divided into four groups based on the TPH levels from low to high: normal (NS11, ND11, NS12, ND12, NS13, ND13), petroleum level 1 (OD5, OD6, OD7, OD8, OD9, OD10, OS8; TPH 0.006–0.018 g/g), petroleum level 2 (OD3, OD4, OS3, OS4, OS7, OS9, OS10; TPH 0.025–0.050 g/g), and petroleum level 3 (OD1, OD2, OS1, OS2, OS5, OS6; TPH 0.108–0.4303 g/g).

Table 1 Geographical, physical, and chemical properties of soils

Sample	TPH (g/g)	WC (g/g)	pH	EC ($\mu\text{s}/\text{cm}$)	TOC (mg/g)	TN (mg/g)	TP (mg/g)	K (mg/g)	Locations (N, E)
OS1	0.4043	0.1428	7.6	2540	160.990	1.024	0.0835	2.9651	38°42'5", 117°29'42"
OD1	0.1080	0.2070	8.0	4030	38.528	2.125	0.1073	1.2158	
OS2	0.3200	0.1548	7.8	1310	106.640	1.431	0.0862	2.6721	38°42'3", 117°29'41"
OD2	0.1884	0.1472	8.3	1810	79.120	1.213	0.1066	2.7725	
OS3	0.0500	0.2060	8.4	735	34.400	2.007	0.0634	1.1206	38°42'14", 117°29'1"
OD3	0.0360	0.2018	8.9	782	20.640	1.179	0.0954	1.0075	
OS4	0.0460	0.2094	8.8	830	34.400	1.744	0.0417	1.0423	38°41'23", 117°30'11"
OD4	0.0250	0.2292	8.6	737	55.040	1.413	0.0415	0.8881	
OS5	0.2520	0.1128	8.1	3660	172.000	2.222	0.1425	1.4106	38°41'11", 117°30'2"
OD5	0.0060	0.2400	8.4	3390	34.400	1.773	0.2183	1.4091	
OS6	0.2200	0.2464	8.2	2880	151.360	2.027	0.1023	1.3164	38°41'31", 117°30'15"
OD6	0.0140	0.2000	8.6	3120	110.080	1.791	0.1005	1.3645	
OS7	0.0400	0.2320	8.3	1930	24.080	1.758	0.1725	0.9475	38°41'30", 117°30'22"
OD7	0.0090	0.2162	8.6	2200	20.640	1.432	0.4073	1.0231	
OS8	0.0160	0.1804	8.1	6640	27.520	1.495	0.4272	1.4935	38°39'24", 117°30'16"
OD8	0.0090	0.2420	8.1	5130	41.280	2.001	0.1032	1.4735	
OS9	0.0450	0.1695	7.8	8150	17.200	2.127	0.1454	1.4632	38°39'13", 117°30'24"
OD9	0.0180	0.1702	8.2	6200	68.800	1.693	0.1425	1.4753	
OS10	0.0260	0.1906	8.3	6710	31.520	1.795	0.5272	1.4926	38°39'23", 117°30'19"
OD10	0.0110	0.2620	8.0	4230	43.280	1.870	0.1132	1.4235	
NS11	0.0000	0.2992	8.2	2902	17.520	1.263	0.2866	2.7705	38°39'18", 117°30'16"
ND11	0.0000	0.1688	8.4	3040	34.400	1.124	0.1315	2.8013	
NS12	0.0000	0.4070	8.1	5000	34.400	3.163	0.0830	1.3123	38°39'25", 117°30'26"
ND12	0.0000	0.3440	8.6	2500	6.880	1.856	0.0640	1.3792	
NS13	0.0000	0.1789	8.5	1792	25.220	1.233	0.1876	2.7455	38°40'25", 117°29'3"
ND13	0.0000	0.1708	8.3	3230	36.400	1.234	0.1235	2.7913	

OS, oil-contaminated surface soils; OD, oil-contaminated deep soils; NS, nonpetroleum-contaminated surface soils; ND, nonpetroleum-contaminated deep soils; TPH, total petroleum hydrocarbon; WC, water content; EC, electrical conductivity; TP, total phosphorus; TN, total nitrogen; K, potassium; TOC, total organic carbon

TPH was significantly negatively correlated with water content and significantly positively correlated with potassium (Supplementary Table S1). WC was significantly negatively correlated with potassium. The pH value was significantly negatively correlated with EC and potassium. EC was significantly positively correlated with TN, TP, and potassium. TN was significantly negatively correlated with potassium.

Statistics for 16S rRNA Sequencing Data

After sequencing the V4 region of the 16S rRNA of 26 soil samples, the number of total tags was 1,737,904, and the average number of tags per sample was 66,842. A total of 81,548 OTUs were obtained through OTU cluster analysis (Supplementary Fig. S1), and the average number of OTUs per sample was 3136. OS7 (5060) had the highest number of OTUs, and OD10 (804) had the lowest.

For OTU picking, we obtained 9724 OTUs from 26 soil samples. The sum of OTUs in oil-contaminated soils (OS: 7389; OD: 7534) was higher than that of nonpetroleum-contaminated soils (NS: 5209; ND: 5007) (Fig. 1). Moreover, shared OTUs took up more than 67% between NS and ND (3497/5209) and OS and OD (6035/7389). And shared OTUs took up about 56% between NS and OS

(4268/7389) and ND and OD (4256/7534). The number of OTUs shared by all four groups was 3146. However, 775 and 948 species of bacteria were unique to the oil-contaminated surface and deep soils, respectively. These values demonstrate that the oil-polluted environment had produced a large number of new OTUs, and these bacteria are likely to use petroleum hydrocarbons as a carbon source.

Bacterial Community Composition and Alpha Diversity Analysis

At the phylum level (Fig. 2), *Proteobacteria* (30.3–54.2%), *Firmicutes* (6.3–16.9%), and *Bacteroidetes* (9.2–14.8%) were the three most abundant species found in NS. *Proteobacteria* (8.0–86.6%), *Cyanobacteria* (24.5–25.7%), and *Firmicutes* (6.7–61.6%) were the three most abundant species found in ND. *Proteobacteria* (34.8–61.9%), *Firmicutes* (4.6–29.7%), and *Bacteroidetes* (4.0–13.8%) were the three most abundant species found in OS. *Proteobacteria* (8.0–86.3%), *Firmicutes* (6.6–61.2%), and *Actinobacteria* (0.1–15.7%) were the three most abundant species found in OD.

ANOSIM results showed that there were significant differences in bacterial communities between the NS and ND, NS and OS, and ND and OD. It also showed that there were no significant differences in bacterial communities between OS and OD (Supplementary Table S2).

Observed species were more numerous, and bacterial diversity was richer in the surface soil samples than in the deep soil samples for both nonpetroleum- and oil-contaminated soils. In addition, normal surface soils had the most diverse evolutionary lineages (Supplementary Table S3).

Rarefaction curves of the number of observed species number in most soils gradually plateaued. However, in OS7, OD9, OS2, and OD4, the slope of the rarefaction curves was still steep when the sequencing data volume reached 40,000, indicating that if the sampling continued, new OTUs would be detected (Supplementary Fig. S2). However, as shown in Fig. S3, the end of each sample curve tended to be straight, indicating that the diversity of microbial communities would not increase. These results indicate that the number of sequencing samples can sufficiently predict microbial diversity (Supplementary Fig. S3).

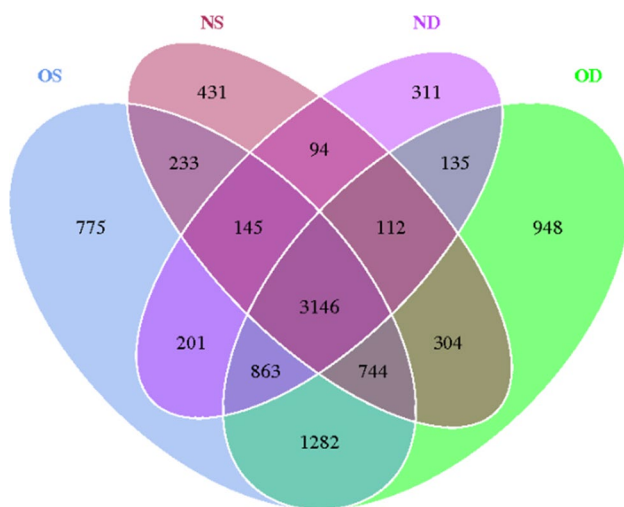
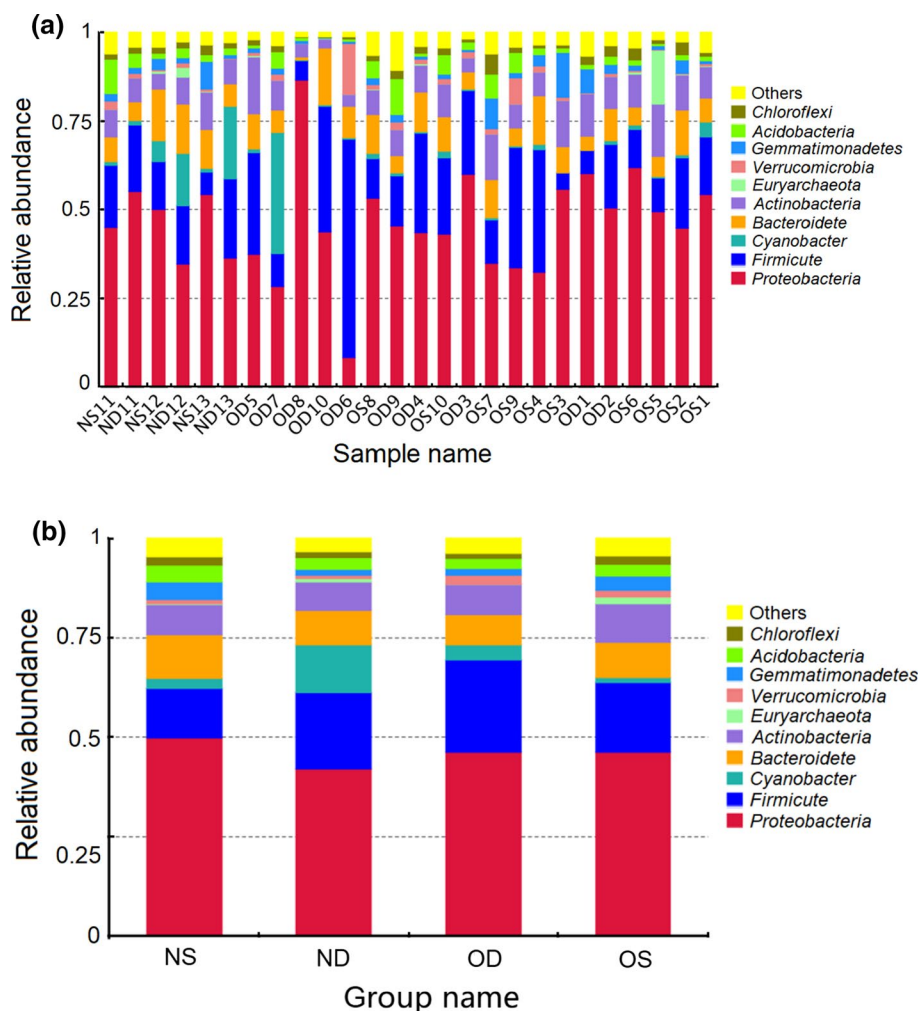


Fig. 1 Venn diagrams of OTUs included in four groups of soils

Fig. 2 Column accumulation chart of species abundance (phylum level). **a** From left to right, TPH gradually increases across 26 samples and **b** four group samples



Correlations Between Environmental Factors and the Bacterial Community

TPH, TOC, and TN had the greatest impacts on the soil bacterial community in the Dagang Oilfield; pH had a significant impact on surface soils but no significant impact on deep soils. TP had a significant impact on deep soils but no significant impact on surface soils. However, TN had a significant impact on all soils. WC, K, and EC had no significant impacts on the soil bacterial community in the Dagang Oilfield (Table 2).

TPH was significantly negatively correlated with the bacterial diversity and the number of bacterial species; pH was significantly negatively correlated with bacterial diversity. TOC was significantly negatively correlated with the number of bacteria and the number of observed bacterial species. TN was significantly negatively correlated with the number of observed bacterial species. TP was significantly negatively correlated with sample coverage (Supplementary Fig. S4).

The two main axes explained 44.29% of the soil bacteria diversity variations. The soil bacterial communities were clustered by color, indicating that the TPH determines the

soil community (Fig. 3). In surface soils, the most important environmental factor affecting the distribution of bacterial communities was TOC, followed by TPH and TN. In deep soils, the most important environmental factors affecting the distribution of bacterial communities were TPH, followed by TOC and TN. In summary, both CCA and the mental test indicated that TPH, TOC, and TN had the greatest impact on the soil bacterial community in the Dagang Oilfield.

Effects of Petroleum Pollution Concentration on the Bacterial Community

The OTUs with the largest abundance differences in soil bacteria were listed in the LEfSe analysis plot. *Cyanobacteria*, *Rickettsiales*, *Sphingomonadales*, *Thermoleophilia*, *Rhizobiaceae*, *Sphingomonadaceae*, *Bradyrhizobiaceae*, *Solirubrobacterales*, *Bradyrhizobium elkanii*, *Bradyrhizobium*, *Galettalles*, and *Methylobacteriaceae* were suitable to be grown in nonpetroleum-contaminated soils. *Acidobacteria*, *Bacteria*, *Acidobacteriales*, *Faecalibaculum*, *Bryobacter*, *Rhizomicrobium*, *Acidibacter*, *Candidatus Solibacter*,

Table 2 Effects of environmental factors on bacterial abundance in soil

Variable	Values	All	Surface	Deep
TPH	<i>r</i>	0.33	0.3	0.44
	<i>p</i>	0.02	0.03	0.02
TOC	<i>r</i>	0.36	0.35	0.37
	<i>p</i>	0.01	0.01	0.03
pH	<i>r</i>	0.18	0.25	0.17
	<i>p</i>	0.18	0.04	0.19
WC	<i>r</i>	0.05	0.02	0.09
	<i>p</i>	0.15	0.17	0.13
TN	<i>r</i>	0.24	0.28	0.32
	<i>p</i>	0.02	0.02	0.04
TP	<i>r</i>	0.21	0.18	0.31
	<i>p</i>	0.01	0.06	0.03
K	<i>r</i>	0.03	0.04	0.02
	<i>p</i>	0.55	0.53	0.36
EC	<i>r</i>	0.07	0.09	0.06
	<i>p</i>	0.18	0.19	0.16

Surface: surface soils; Deep: deep soils; and All: all soils. The effect increases as the *r* value increases. *p* < 0.05 represents a significant correlation

Sphingobacteriaceae, *Granulicella*, and *Halanaerobiales Mucilaginitacter* were suitable to be grown in oil-contaminated soils (TPH 0.0250–0.0500 g/g). *Acidobacteria*, *Bacteria*, *Acidobacteriales*, *Faecalibaculum*, *Bryobacter*, *Rhizomicrobium*, *Acidibacter*, *Candidatus Solibacter*, *Sphingobacteriaceae*, *Granulicella*, *Halanaerobiales*, and *Mucilaginitacter* were suitable to be grown in oil-contaminated soils (TPH 0.108–0.4043 g/g) (Supplementary Fig. S5).

Most samples could be clustered together according to the TPH content. *Phyllobacterium* and *Ralstonia* were widely distributed in soils (except for OD1, OS3, and OS5), indicating that they were the predominant bacteria in the Dagang Oilfield. Therefore, they had a powerful ability to adapt to the environment. *Streptococcus* and *Pontibacter* were the two most abundant genera in nonpetroleum-contaminated surface soils. *Unidentified-Chloroplast* and *unidentified-Mitochondria* are the two most abundant genera in nonpetroleum-contaminated deep soils (Supplementary Fig. S6). *Marinobacter*, *Alcanivorax*, and *Mycobacterium* were the three most abundant genera in oil-contaminated soils. In addition, *Alcanivorax*, *Mycobacterium*, *Pseudomonas*, and *Micrococcus* are petroleum hydrocarbon-degrading bacteria. These results show that the soil of Dagang Oilfield has abundant oil-degrading microorganisms.

The two main axes explained 37.26% of the soil bacterial diversity, and soil bacterial community was clustered by both color and size, indicating that TPH, TN, and TOC determine the soil community (Fig. 4).

Soil Bacterial Metabolic Pathway Prediction (Third Level) by KEGG

Oil-contaminated soils had more methane metabolic pathways and more polycyclic aromatic hydrocarbon degradation pathways than normal soils (Fig. 5). Methane metabolic and polycyclic aromatic hydrocarbon degradation pathways use petroleum hydrocarbons as substrates. OS8, OD6, OD5, OS2, OS9, and OD10 had abundant methane metabolic pathways ($Z > 3$). OS5, OD6, OS6, OS10, and OS4 had abundant polycyclic aromatic hydrocarbon degradation pathways ($Z > 3$).

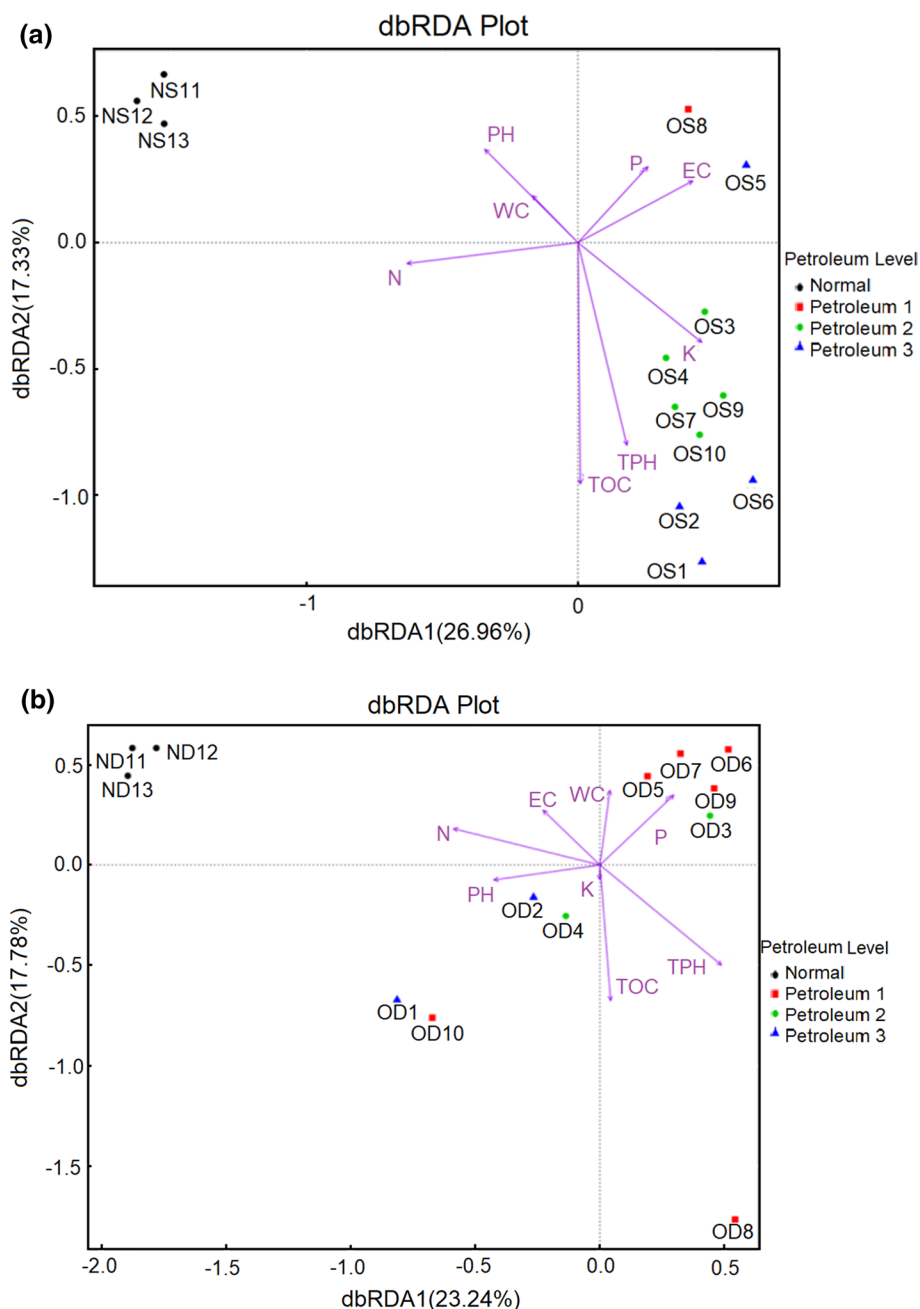
Discussion

In oil-contaminated soils, *Euryarchaeota* was the most abundant phylum of Archae bacteria, which appears to be consistent with the study of Sutton et al. [20]. When TPH was 0.04–0.4043 g/g (except in OS3 and OD1), the relative abundance of *Gammaproteobacteria* was positively correlated with the concentration of oil pollution (Fig. 3), which agrees with a previous report [21]. In oil-contaminated soils, *Proteobacteria* and *Firmicutes* were the dominant bacteria, suggesting that they may have petroleum hydrocarbon metabolic pathways (Fig. 2). Yakimov et al. [22] demonstrated that *Proteobacteria* and *Firmicutes* can use petroleum hydrocarbons as a carbon source. The relative abundance of *Cyanobacteria* in nonpetroleum-contaminated soils was higher than that in oil-contaminated soils for both surface and deep soils, indicating that petroleum pollution had inhibited the growth of *Cyanobacteria* (Fig. 2). However, *Cyanobacteria* can use petroleum hydrocarbons as carbon source in oil-contaminated water [23]. We speculate that the dispersion and accumulation of petroleum in water are different from that in soil. Petroleum spills in water show horizontal diffusion and dilution in water. However, petroleum spills on land show vertical penetration and absorption by humus [24]. This phenomenon leads to the accumulation of volatile poisonous substances in the soil, causing a significant damage to the *Cyanobacteria*.

Proteobacteria, *Chloroflexi*, *Euryarchaeota*, and *Verrucomicrobia* belong to the top 10 most abundant phyla in oil-contaminated soils (Fig. 2), and their metabolic pathways are interconnected. *Chloroflexi* can degrade petroleum hydrocarbons by anaerobic respiration. *Proteobacteria* can consume petroleum hydrocarbons and produce organic acids [25], while *Euryarchaeota* can use these organic acids as intermediates [26]. *Verrucomicrobia* can also convert methane into formaldehyde and carbon dioxide [27].

ANOSIM showed that there were significant bacterial community differences between surface and normal deep soils but no significant bacterial community differences

Fig. 3 Redundancy analysis (RDA) of **a** surface soils and **b** deep soils (The influence of environmental factors on the bacterial community is indicated by arrows. The length of each arrow represents the impact of the corresponding environmental factor on the community distribution. The longer is the length of the arrow, the greater is the effect.)

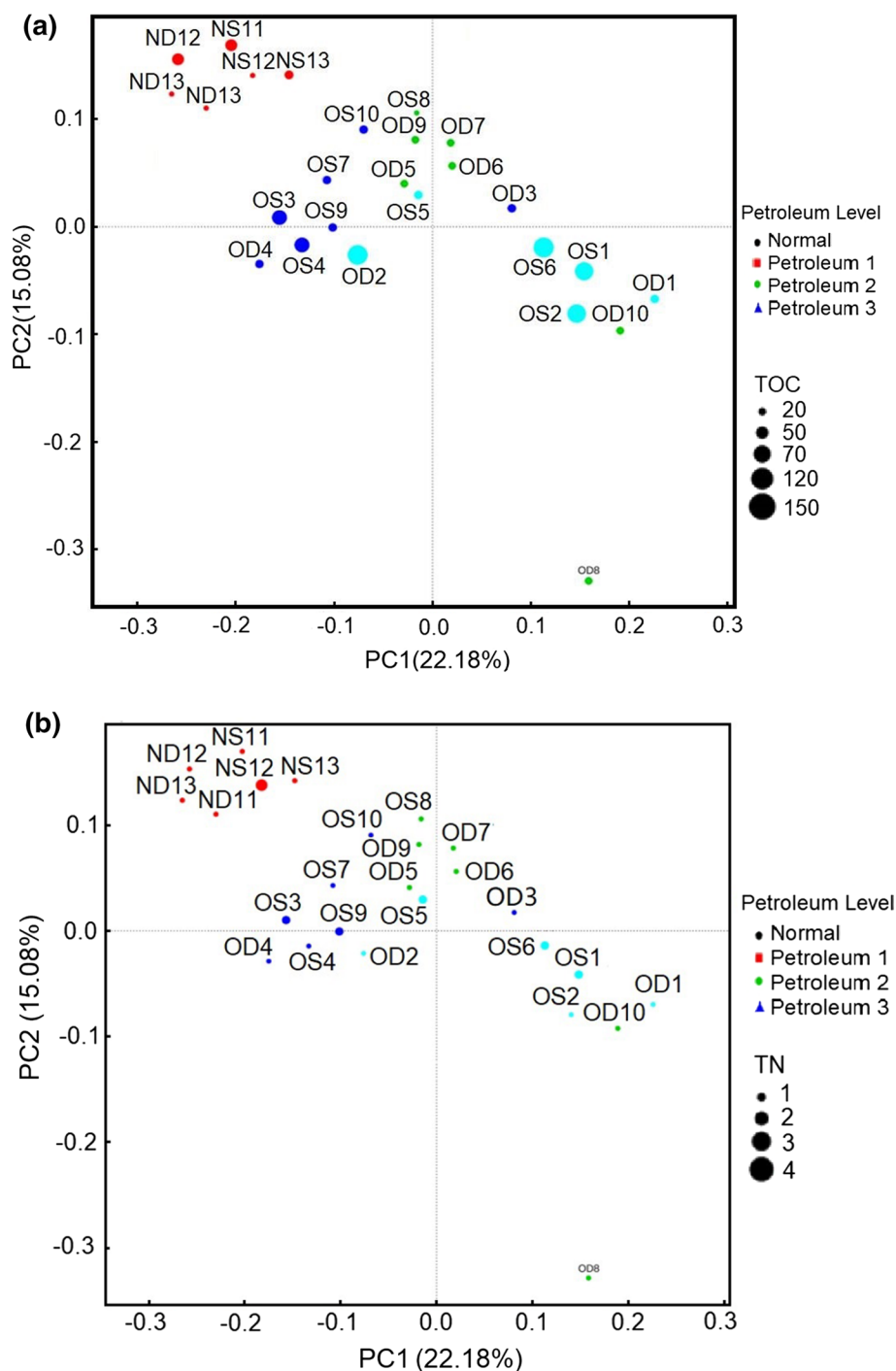


between surface and oil-contaminated deep soils (Supplementary Table S2) [28]. We assume that oxygen content and nutrients in nonpetroleum-contaminated soils are different at different depths. However, in oil-contaminated soils, petroleum pollution is the main factor affecting community diversity. Oil adheres to soil particles and binds them together, which reduces the difference in the soil oxygen content between the surface and deep layers. As a result, the effects of depth on the distribution of bacterial communities in oil-contaminated soils are reduced. We found that most oil-contaminated soils could be clustered based on TPH (heat map, Supplementary Fig. S7).

LefSe [29] has shown that the TPH preference of bacteria has generated a novel way to rehabilitate oil-contaminated soils: the addition of oil-degrading bacteria to oil-contaminated soils [8]. This targeted bioenhancement method not only improves the efficiency of the repair of oil-contaminated soils but also reduces the cost.

Redundancy analysis (RDA) and mental test showed that TPH, TOC, and TN had the greatest impact on the soil bacteria community in the Dagang Oilfield (Table 2 and Fig. 3). The nutrient element that had the greatest impact on bacterial diversity was TN, which is consistent with the findings of Ward et al. [30]. The TPH was

Fig. 4 PCoA of **a** TPH and TOC and **b** TPH and TN



significantly negatively correlated with bacterial diversity (Supplementary Fig. S4). However, there have been varying conclusions about the impact of oil pollution on the diversity of soil bacteria. Li et al. [9] found that the oil content of soils was positively correlated with bacterial diversity in the Daqing and Karamay Oilfields. Gao et al. [13] found that the oil pollution concentration in the oil-contaminated Yellow River Delta area was positively correlated with the soil microbial diversity and evenness

index. However, Nakatsu [31] and Green et al. [32] found that the concentration of PAHS was negatively correlated with the diversity of microorganisms. Wilson and Jones [33] found that an increase in oil content reduced the soil microbial diversity in the Loess Plateau.

Ward et al. [30] discovered that the Amoco oil degradation rate was much higher under aerobic conditions than that in anaerobic conditions. When petroleum hydrocarbons were used as a carbon source, sufficient trace element

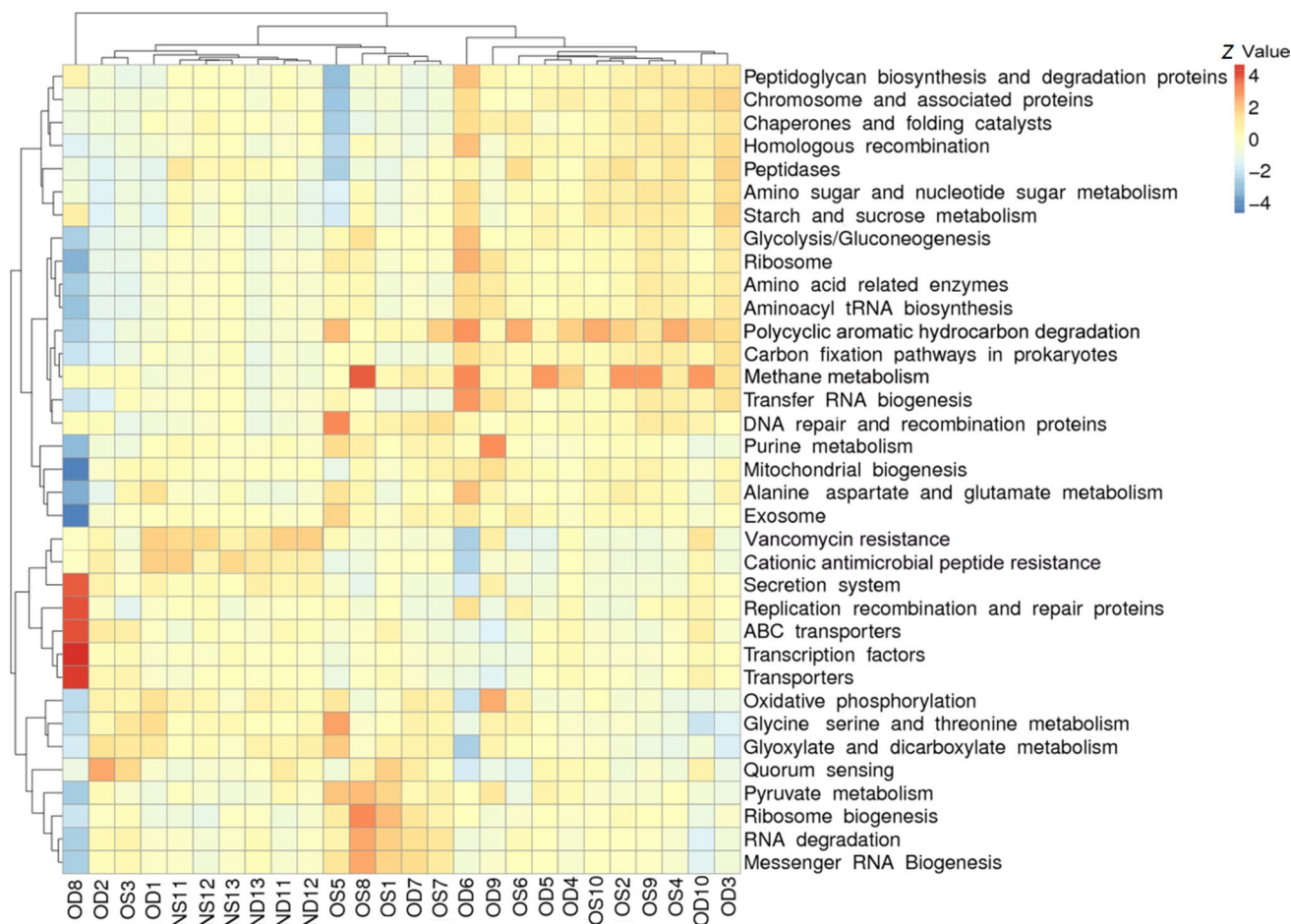


Fig. 5 Bacteria metabolic pathway heat map (third level). (The ordinate is the top 35 most abundant metabolic pathways; the abscissa is a sample clustering tree of the evolutionary relationships of metabolic pathways. Z value: normality of relative abundance.)

content had a decisive influence on the microbial degradation of hydrocarbons. Margesin et al. [34] found that after adding N, P, and other nutrients to oil-contaminated areas, the total number of bacteria increased several fold over a short time and was accompanied by a rapid decrease in the content of *n*-alkanes and isoprenoids. This observation indicates the importance of environmental factors to petroleum-degrading bacteria in the repair of petroleum-contaminated soils. The exploration of bacterial community diversity and its correlation with environmental factors in the oil-contaminated Dagang Oilfield soils has greatly contributed to the use of biostimulation to repair petroleum-contaminated soils.

Conclusions

In this study, we investigated 26 samples from the Dagang Oilfield to clarify the bacteria community distribution along several environmental gradients in both surface (5–10 cm)

and deep (35–40 cm) petroleum-contaminated soils. There were significant differences in the structure of bacterial communities between nonpetroleum-contaminated surface and subsurface soils, but no significant difference was observed in the structure of bacterial communities between petroleum-contaminated surface and deep soils. We also found that total petroleum hydrocarbon, total organic carbon, and TN content were the three main drivers of the bacteria community distribution in the Dagang Oilfield. Our findings can provide references for sampling strategy and bacteria community compositions in petroleum-contaminated soils. Also, the distribution patterns of the bacteria community and taxa, especially some petroleum-degrading bacteria (mainly *Alcanivorax* and *Mycobacterium*), along the environmental gradients in petroleum-contaminated soils provided references for further new gene exploitations, such as some petroleum hydrocarbon degradation genes. Finally, further in-depth investigations on microbial communities in petroleum-contaminated soils in large scale or with greater depth were expected.

Acknowledgements This work was supported by the Major Research Plan of Tianjin (No. 16YFXTSF00460) and the National Natural Science Foundation of China (No. 21878220).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Lee EH, Lee SH, Cho KS (2011) Bacterial diversity dynamics in a long-term petroleum-contaminated soil. *J Environ Sci Health Part A* 46(3):281–290
- Klironomos JN, McCune J, Hart M et al (2000) The influence of arbuscular mycorrhizae on the relationship between plant diversity and productivity. *Ecol Lett* 3(2):137–141
- Liu J, Song XY, Sun RL et al (2014) Petroleum pollution and the microbial community structure in the soil of Shengli Oilfield. *Chin J Appl Ecol* 25(3):850–856 (in Chinese)
- Ahmed OE, Eldesoky AM, El Nady MM (2019) The use of polycyclic aromatic in the assessment of marine life and the impact on petroleum pollution in the Suez Gulf, Egypt. *Pet Sci Technol* 37(12):1400–1409
- Rocha LL, Colares GB, Nogueira VLR et al (2016) Distinct habitats select particular bacterial communities in mangrove sediments. *Int J Microbiol* 2016:1–6
- Hong Q, Dong XJ, He LJ et al (2009) Isolation of a biphenyl-degrading *Bacterium*, *Achromobacter* sp. BP3, and cloning of the *bph* gene cluster. *Int Biodeterior Biodegrad* 63(4):365–370
- Sood N, Lal B (2009) Isolation of a novel yeast strain *Candida digboiensis* TERI ASN6 capable of degrading petroleum hydrocarbons in acidic conditions. *J Environ Manag* 90(5):1728–1736
- Ueno A, Ito Y, Yumoto I et al (2007) Isolation and characterization of bacteria from soil contaminated with diesel oil and the possible use of these in autochthonous bioaugmentation. *World J Microbiol Biotechnol* 23(12):1739–1745
- Li YX, Xu Y, Liu LL et al (2012) Five new amicoumacins isolated from a marine-derived *Bacterium bacillus subtilis*. *Mar Drugs* 10(12):319–328
- Hay ID, Lithgow T (2019) Filamentous phages: masters of a microbial sharing economy. *EMBO Rep* 20(6):e47427
- Villanueva L, Navarrete A, Urmeneta J et al (2007) Analysis of diurnal and vertical microbial diversity of a hypersaline microbial mat. *Arch Microbiol* 188(2):137–146
- Xu HR, Luo YH, Chen C et al (2016) 3D shallow structures in the Baogutu area, Karamay, determined by eikonal tomography of short-period ambient noise surface waves. *J Appl Geophys* 129:101–110
- Gao YC, Wang JN, Guo SH et al (2015) Effects of salinization and crude oil contamination on soil bacterial community structure in the Yellow River Delta region, China. *Appl Soil Ecol* 86:165–173
- Abed RMM, Al-Kindi S, Al-Kharusi S (2015) Diversity of bacterial communities along a petroleum contamination gradient in desert soils. *Microb Ecol* 69(1):95–105
- He LF, Xiang XY, Zhang RP et al (2011) Simultaneous determination of polycyclic aromatic hydrocarbons in soil by way of supersonic extraction and synchronous fluorescence spectrometry. *J Saf Environ* 11(5):103–108 (in Chinese)
- Zheng J, Feng JQ, Zhou L et al (2018) Characterization of bacterial composition and diversity in a long-term petroleum contaminated soil and isolation of high-efficiency alkane-degrading strains using an improved medium. *World J Microbiol Biotechnol* 34(2):34
- Wang X, Han Z, Bai Z et al (2011) Archaeal community structure along a gradient of petroleum contamination in saline-alkali soil. *J Environ Sci* 23(11):1858–1864
- Jones CM, Thies JE (2007) Soil microbial community analysis using two-dimensional polyacrylamide gel electrophoresis of the bacterial ribosomal internal transcribed spacer regions. *J Microbiol Methods* 69(2):256–267
- Choe M, Hong SJ, Lim JH et al (2013) Korean paddy soil microbial community analysis method using denaturing gradient gel electrophoresis. *J Appl Biol Chem* 56(2):95–100
- Sutton NB, Maphosa F, Morillo JA et al (2013) Impact of long-term diesel contamination on soil microbial community structure. *Appl Environ Microbiol* 79(2):619–630
- Hazen TC, Dubinsky EA, de Santis TZ et al (2010) Deep-sea oil plume enriches indigenous oil-degrading bacteria. *Science* 330(6001):204–208
- Yakimov MM, Timmis KN, Golyshin PN (2007) Obligate oil-degrading marine bacteria. *Curr Opin Biotechnol* 18(3):257–266
- Head IM, Jones DM, Röling WFM (2006) Marine microorganisms make a meal of oil. *Nat Rev Microbiol* 4(3):173–182
- Icoz I, Stotzky G (2008) Fate and effects of insect-resistant crops in soil ecosystems. *Soil Biol Biochem* 40(3):559–586
- Leahy JG, Colwell RR (1990) Microbial degradation of hydrocarbons in the environment. *Microbiol Rev* 54(3):305–315
- Xia WJ, Shen WJ, Yu L et al (2016) Conversion of petroleum to methane by the indigenous methanogenic consortia for oil recovery in heavy oil reservoir. *Appl Energy* 171:646–655
- Dunfield PF, Yuryev A, Senin P et al (2007) Methane oxidation by an extremely acidophilic bacterium of the phylum Verrucomicrobia. *Nature* 450(7171):879–882
- da Silva DKA, Coutinho FP, Escobar IEC et al (2015) The community of arbuscular mycorrhizal fungi in natural and revegetated coastal areas (Atlantic Forest) in northeastern Brazil. *Biodivers Conserv* 24:2213–2226
- Fang D, Zhao G, Xu X et al (2018) Microbial community structures and functions of wastewater treatment systems in plateau and cold regions. *Bioresour Technol* 249:684–693
- Ward DM, Atlas RM, Boehm PD et al (1980) Microbial biodegradation and chemical evolution of oil from the amoco spill. *Ambio* 9(6):277–283
- Nakatsu CH (2007) Soil microbial community analysis using denaturing gradient gel electrophoresis. *Soil Sci Soc Am J* 71(2):562–571
- Green SJ, Leigh MB, Neufeld JD (2015) Denaturing gradient gel electrophoresis (DGGE) for microbial community analysis. In: McGenity TJ, Timmis KN, Nogales B (eds) *Hydrocarbon and lipid microbiology protocols*. Springer, Berlin, pp 77–99
- Wilson SC, Jones KC (1993) Bioremediation of soil contaminated with polynuclear aromatic hydrocarbons (PAHs): a review. *Environ Pollut* 81(3):229–249
- Margesin R, Hämmerle M, Tschërko D (2007) Microbial activity and community composition during bioremediation of diesel-oil-contaminated soil: effects of hydrocarbon concentration, fertilizers, and incubation time. *Microb Ecol* 53(2):259–269



Wenyu Lu is an Associate Professor in School of Chemical Engineering and Technology, Tianjin University (TJU). He received his Ph.D. in Fermentation Engineering at Tianjin University of Science & Technology. He had his post-doctor training in Biochemistry at TJU and joined TJU as an Associate Professor in 2007. From April 2014 to April 2015, he worked as a visiting scholar in the research of synthetic biotechnology in the Department of Chemical Engineering at the University of

Minnesota. His research primarily focuses on Synthetic Biology and Biochemical Reaction Engineering to achieve two major goals: (1) elucidating biosynthesis process and intracellular production based on dynamic regulation and (2) strengthening microbial fermentation process based on the overall productivity. His research team work on gene elements mining, module construction, biosynthesis pathway regulation and mechanism illustrating, microbial cell factories construction, and microbial cells production process enhancing and scale-up, covering the process of fundamental research and application.