



典型气田的指示微生物异常分子诊断：以杭锦旗气田为例

刘海峰¹，许科伟^{2*}，颜正飞¹，顾磊²，郑旭莹²，张凤山³，宿玲恰^{1*}

1 江南大学生物工程学院，江苏 无锡 214122

2 中国石油化工股份有限公司石油勘探开发研究院无锡石油地质研究所，江苏 无锡 214126

3 山东华泰纸业股份有限公司&山东黄三角产业技术研究院 山东 东营 257335

刘海峰，许科伟，颜正飞，顾磊，郑旭莹，张凤山，宿玲恰. 典型气田的指示微生物异常分子诊断：以杭锦旗气田为例[J]. 微生物学报, 2024, 64(6): 2008-2024.

LIU Haifeng, XU Kewei, YAN Zhengfei, GU Lei, ZHENG Xuying, ZHANG Fengshan, SU Lingqia. Molecular diagnosis of indicator microorganisms in typical gas fields: a case study of Hangjinqi Gas Field[J]. Acta Microbiologica Sinica, 2024, 64(6): 2008-2024.

摘要：【目的】油气微生物勘探技术因其多解性小、信噪比高、受环境影响小和经济快速的特点，日益受到勘探者的重视。然而，多数微生物勘探应用是基于非原位的实验室培养分析，不能准确而全面反映漫长地质历史过程中油气资源微生物的原位动态变化规律。本研究以杭锦旗气田为例，分析气区上方和背景区的微生物群落结构和发育特征差异，识别地表油气微生物异常。【方法】对采集自杭锦旗的新昭和什股壕大营圈闭土壤样本中细菌 16S rRNA 基因进行测序，结合理化参数，比较微生物多样性差异，分析理化参数对微生物分布的影响，识别微生物异常。通过共生网络分析，探索气藏上方地表土壤中微生物群落构建过程和功能组成。【结果】群落组成分析发现在杭锦旗地区，放线菌门和变形菌门是主要的微生物类群，占菌落总体丰度的 72.47%，理化参数相关性分析表明该地区的微生物分布与理化参数无明显相关关系。群落多样性分析显示，气区和背景区的群落结构存在显著差异。共生网络分析发现气区微生物群落呈现出非随机性和连通性，表明确定性因素在微生物群落构建中占主导作用。模块化的共发生网络揭示微生物群落形成了特定的功能模块，不同模块的节点可能具有不同的功能。【结论】通过对杭锦旗气区和背景区的微生物多样性分析发现了新昭和什股壕气区的指示菌属。根据对杭锦旗气区的共生网络分析，发现关键指示类群为出芽单胞菌属、土壤红色杆形菌属、假诺卡氏菌属、短芽孢杆菌属、气微菌

资助项目：江苏省重点研发计划(BE2021625)

This work was supported by the Key Research and Development Program of Jiangsu Province (BE2021625).

*Corresponding authors. E-mail: SU Lingqia, sulingqia@jiangnan.edu.cn; XU Kewei, xukw.syky@sinopec.com

Received: 2023-10-09; Accepted: 2023-12-15; Published online: 2023-12-21

属和类诺卡氏菌属, 主要的功能模块包括碳氮循环和有机物质降解模块, 共同参与气田上方土壤中烃类物质的降解。

关键词: 微生物勘探; 高通量测序; 分子诊断; 共发生网络; 杭锦旗气田

Molecular diagnosis of indicator microorganisms in typical gas fields: a case study of Hangjinqi Gas Field

LIU Haifeng¹, XU Kewei^{2*}, YAN Zhengfei¹, GU Lei², ZHENG Xuying², ZHANG Fengshan³, SU Lingqia^{1*}

1 School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China

2 Wuxi Research Institute of Petroleum Geology, Sinopec Petroleum Exploration and Production Research Institute, Wuxi 214126, Jiangsu, China

3 Shandong Huatai Paper Co., Ltd. & Shandong Yellow River Delta Research Institute Co., Ltd., Dongying 257335, Shandong, China

Abstract: [Objective] Microbial prospecting for oil and gas, characterized by high resolution, high signal-to-noise ratio, minimal environmental interference, low costs, and short time consumption, garners increasing attention from exploration experts. However, in most cases, microbial prospecting is based on laboratory culture and analysis, which cannot accurately and comprehensively reflect the *in-situ* dynamic changes of microbiota in oil and gas resources in the geological history. In this study, we compared the microbial community structure and developmental characteristics between the gas-producing zone and the background zone in Hangjinqi Gas Field, aiming to identify the surface microbial anomalies related to oil and gas. **[Methods]** We conducted the bacterial 16S rRNA gene sequencing for the soil samples collected from Xinzaoyao and Shiguohao areas of Hangjinqi. Furthermore, we compared the microbial diversity, analyzed the impacts of physicochemical parameters on microbial distribution, and identified microbial anomalies. The co-occurrence network analysis was employed to explore the assembly process and functional composition of microbial community in the surface soil above the reservoir. **[Results]** In the Hangjinqi area, *Actinobacteria* and *Proteobacteria* were dominant, accounting for 72.47% of the total microbial abundance. The correlation analysis of environmental factors with microbial abundance showed that the distribution of microorganisms in this area was not significantly correlated with environmental factors. The microbial community structure presented significant differences between the gas-producing area and the background area. The co-occurrence network analysis of the gas-producing area revealed non-randomness and connectivity in the microbial community, indicating deterministic factors play a dominant role in the construction of microbial communities. Modular co-occurrence network analysis revealed the formation of specific functional modules within the microbial community, and different modules possibly served different functions. **[Conclusion]** By comparing the microbial diversity between the gas-producing and background area of Hangjinqi area, we identified the indicator genera in the

gas-producing fields of Xinzaoyao and Shiguohao. Furthermore, the co-occurrence network analysis identified *Gemmatimonas*, *Solirubrobacter*, *Pseudonocardia*, *Brevibacillus*, *Aeromicrobium*, and *Nocardioideae* as the key taxa in the gas-producing area, which were associated with the main functional modules of carbon and nitrogen cycling and organic matter degradation, contributing to the degradation of hydrocarbons in the surface soil of the gas-producing area.

Keywords: microbial prospecting; high-throughput sequencing; molecular diagnosis; co-occurrence network; Hangjinqi Gas Field

油气微生物勘探技术是一项通过检测近地表土壤的微生物异常来预测下伏油气藏的存在、分布和性质的技术。其基本原理是：在油气储层压力的驱动下，地下深部油气藏的轻烃组分(C₁-C₅)持续地向地表作垂直扩散和迁移，这种轻烃的微渗漏现象在近地表土壤中引发了以轻烃气为唯一能量来源的微生物的发育并形成微生物异常^[1-2]。这项技术作为油气勘探方法的一个分支，由于其具有直接、有效、快速、经济和多解性少等优势，越来越受到全球油气勘探公司的重视^[3-4]。传统的油气微生物勘探技术是通过检测可培养油气指示微生物的数量反映对应区块的油气微生物异常^[5-7]。然而，大量的油气指示微生物无法通过纯培养技术获得，这无法具体地表征某一区块的油气指示微生物的发育特征与群落结构。随着高通量测序技术、生物信息学技术的快速发展，原位分析近地表土壤的微生物群落结构与异常成为可能^[8]。

然而，鉴于高通量测序所产生的基因组数据集规模庞大且结构复杂，提取出其中有意义的信息和模式是一项艰巨的任务^[9]。由于运移至地表的烃类浓度很低，油气指示微生物并不是土壤群落中的优势菌群。如何从这些基因组数据集中有效提取关键种属的多样性信息指导油气勘探至关重要。目前，已有利用高通量测序对木里地区水合物轻烃微渗漏微生物群落的研究，但该研究重点聚焦于微生物群落的多样性，尚未对油气藏上方地表土壤的非优势指示菌属进行分析^[10]。线性

判别分析效应大小(linear discriminant analysis effect size, LEfSe)是一种发现和解释高纬度数据生物标识(分类单元、通路、基因)的分析工具，可以实现两个或者多个分组之间的比较，同时也可进行分组内部亚组之间的比较分析，有利于找到气区与背景区之间存在显著差异的非优势菌属。目前大多数分析技术都聚焦于群落单一特性的研究，很少通过序列数据探索环境样品中共存的微生物类群之间的直接或间接相互作用。通过记录复杂多样群落中类群之间的相互作用(即共生网络分析)有助于揭示气区微生物群落结构的构建过程，确定气区未培养指示微生物所占据的功能角色或环境生态位。

本研究以杭锦旗气田为例，选取新召和什股壕大营圈闭及其油气背景区上方土壤样本进行16S rRNA基因高通量测序，结合共发生网络分析，以期揭示气田上方微生物的分布规律并鉴定主导性微生物群体，以强调不同气区间指示微生物的差异特征。此外，本研究还确定了气区构建的主导因素，明晰了关键物种在其中的作用，揭示了主要功能模块的关联网络。

1 材料与方法

1.1 样品采集及理化参数测定

杭锦旗地区位于鄂尔多斯盆地北部，由于该区自古生代以来一直处于相对隆起状态，具有较大的天然气勘探潜力^[11]。然而该区的砂体储层表现出了高度的非均质性和复杂性，其中储集砂体

的厚度变化很大且空间分布也很复杂。这些砂体主要以条带状的形式分布，横向变化速度很快，是典型的低孔低渗储层。这种特殊的储层性质增加了天然气富集区的预测难度，为油气勘探和评价带来了很大的挑战^[12]。

由于杭锦旗地区砂体分布不规则，通过一般的网格布点原则很难明确具体的气区背景区位置。通过油公司现场实际的钻井产气情况进行取样，选取其中产气井位置作为明确的气区样品，而干井位置采集为背景区样品。根据五点采样法，选取钻井点周围 4 个距离相等的点，提取 5 个样本点的 DNA 并混合作为 1 个样本。在杭锦旗地区的新召和什股壕大营圈闭分别采集气区样品(8 个)和背景区(8 个)样品，共计 32 个样品，对这些样

品进行编号分组(采样位置见图 1)。钻井位置根据前期的大量基础性工作得到的数据进行选取^[13]，土壤采样严格按照《土壤环境监测技术规范》(<https://www.mee.gov.cn/image20010518/5406.pdf>)进行。野外采样时，采集土壤表层 80 cm 深度样品，并且尽量避免草根、石块对样品的影响。每个样品采集 100 g 土样装袋密封，采样结束后运输回实验室-20 °C 保存。

样本中土壤含水量测定：将铝盒 105 °C 烘干 2 h 后称重，精确称取土壤 5 g，加入铝盒中 105 °C 烘干 6 h，烘干土样放入干燥器冷却 20 min 后立即称重。含水量(%)=(土样湿重-土样干重)/土样湿重×100。土壤 pH 值的测定：直接取 5 g 过 2 mm 筛的土壤按 1:2.5 的比例加入去离子水，振荡混

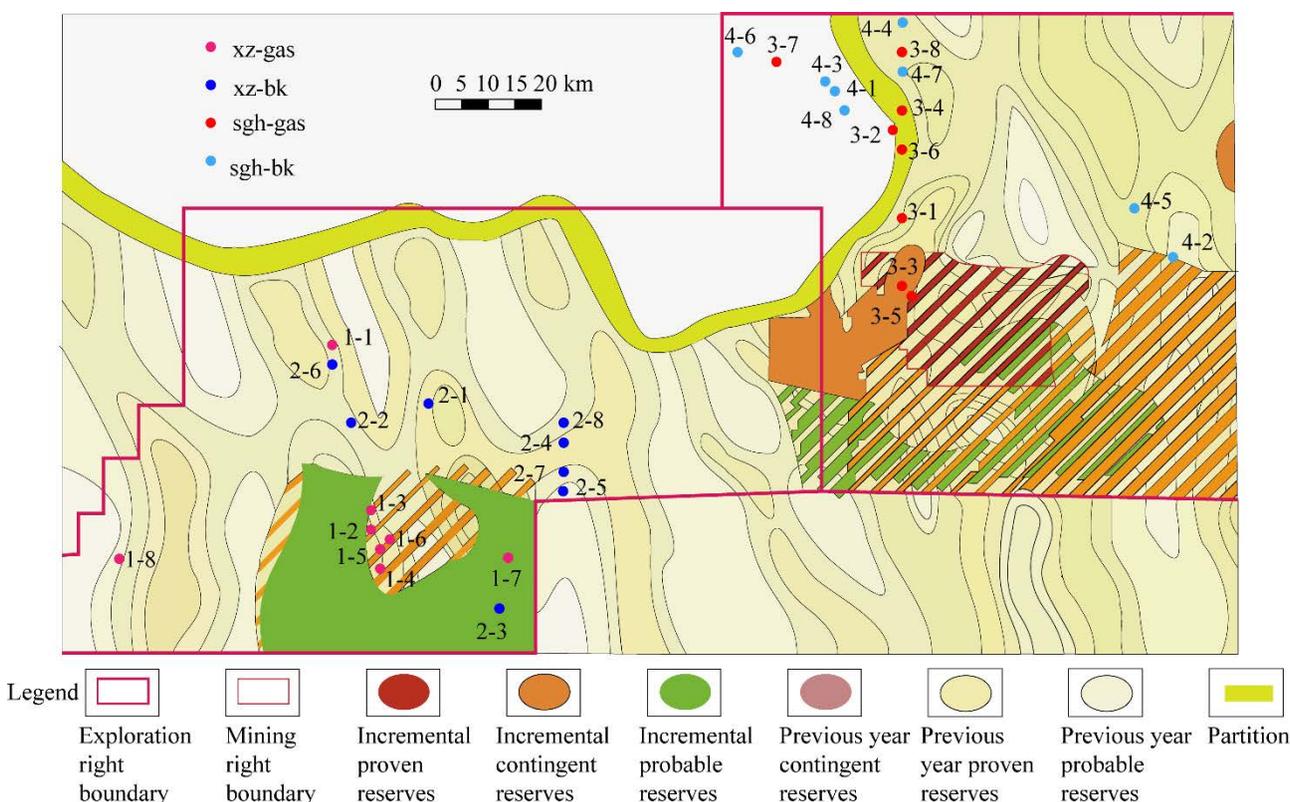


图 1 采样位置图

Figure 1 Map of sampling locations. xz-gas represents samples locating in Xinzaog gas field, xz-bk represents samples locating in Xinzaog background field, sgh-gas represents samples locating in Shiguohao gas field, sgh-bk represents samples locating in Shiguohao background field. The symbols in the image below represent the same meaning.

匀后,用精密 pH 计测定上清的 pH 值。土壤溶解态氮盐测定:称取新鲜土壤 1–4 g 置于 50 mL 离心管中,加入 20 mL 2 mol/L 的 KCl 浸提液,150 r/min 振荡 30 min,静置 0.5 h 后,滤纸过滤上清至 50 mL 离心管,流动分析仪测定滤液中的 NH_4^+ 、 NO_3^- 。土壤 SO_4^- 、 Cl^- 离子测定:直接称取 5 g 土壤 105 °C 烘干至恒重后,研磨破碎,过 100 目筛,取 10 g 加入无 CO_2 超纯水 50 mL,振荡 3 min,之后 5 000 r/min 离心 5 min,上清液用滤纸过滤,将滤液 10 000 r/min 离心 10 min,上清液过 0.22 μm 滤膜后由离子色谱仪(赛默飞世尔科技公司)测定 SO_4^- 、 Cl^- 离子浓度。

1.2 DNA 提取和纯化

称取 5 g 土壤样品,利用离心沉淀法提取土壤中的菌落。采用 TIANamp Bacteria DNA Kit [天根生化科技(北京)有限公司]提取土壤微生物总 DNA。采用 V3–V4 区通用引物对 338f (5'-ACTCCTACGGGAGGCAGCAG-3') 和 806r (5'-GGACTACHVGGGTWTCTAAT-3') 对检测合格的环境微生物 DNA 样品的指定区域进行 PCR 扩增、文库制备、文库质检和定量,使用设定的 TAG 序列进行样本区分。PCR 反应体系(25 μL): dNTPs 0.5 μL , *Taq* 酶(5 U/ μL) 0.25 μL , 10 \times PCR 缓冲液 2.5 μL , DNA 1.0 μL , 正、反向引物(10 $\mu\text{mol/L}$)各 1.0 μL , ddH₂O 18.75 μL 。PCR 反应条件:94 °C 预热 3 min;94 °C 变性 1 min,55 °C 复性 30 s,72 °C 延伸 30 s,30 次循环;72 °C 延伸 10 min,4 °C 保存。采用 Illumina HiSeq 2500 高通量测序平台对检测合格的文库进行测序。

1.3 16S rRNA 扩增子的序列分析

根据 barcode 序列区分样本,将提取出的数据以标准的 Fastq 格式保存。使用 Fastq 文件对测序样品进行数据质量评估。评估后根据配对端 (paired end, PE) reads 之间的重叠 (overlap) 采用 Flash 软件对数据进行拼接,对低质量的数据进

行过滤^[14]。根据 97% 的序列相似性水平,利用 QIIME 2 软件包中的 uclust 方法进行分类操作单元 (operational taxonomic unit, OTU) 聚类分析,在聚类过程中,采用 QIIME 2 软件中的 Usearch 8.0 软件进行嵌合体序列的检测及过滤,得到 OTU 的代表序列,并生成物种组成表 (OTUs 表);然后基于 Silva 参考数据库,采用 RDP 分类学注释 (RDP classifier) 贝叶斯算法对 97% 相似水平的 OTU 代表序列进行物种分类学 (taxonomy) 注释。

1.4 数据分析

基于得到的 OTUs 表,在 R 中进行 α 多样性和 β 多样性分析。样本的 α 多样性通过 Shannon 指数和 Simpson 指数评估。利用 R 中的 vegan 包计算 unweighted-unifrac 距离矩阵,同时进行主坐标分析 (principal coordinates analysis, PCoA) 评估。根据测定的环境因子,基于主坐标分析的第一轴作为微生物分布的评估,利用 Pearson 及 Spearman 两种相关性分析方法,评估环境因子对微生物分布的影响。

在比较不同分组之间的差异时,采用 wilcox 秩和检验得到不同分组间种水平的差异分析显著性 P 值,并过滤掉 P 值小于 0.05 的物种。利用这些差异分析显著性 P 值绘制曼哈顿图,在比较产气区样品和背景区样品的物种差异时,视产气区样品物种平均丰度大于 0.03% 为显著富集,而在比较不同地区的产气区样品物种差异时,视地区样品物种平均丰度较大且大于 0.03% 为在该地区显著富集。同时,使用 python 中的 pandas 包对 OTUs 表格进行预处理,使其符合 LEfSe 在线分析的文件格式。

通过对分类单位属水平的丰度进行两两比较,识别出的所有呈正或负相关性的属构成了一个相关网络,其中每个节点代表一个属,每条边表示节点之间的强相关关系。为了描述所得网络的拓扑特征,使用 R 环境中的 igraph 包计算了

一组指标(节点数、边数、平均路径长度、网络直径、平均度、图密度、聚类系数和模块性), 并使用交互式平台 Gephi 对网络进行可视化。同时, 生成了 1 000 个 Erdos-Renyi 随机网络, 用于与真实网络的拓扑进行比较, 其中每条边被分配给任意节点的概率相等。

2 结果与分析

2.1 理化参数测定和微生物多样性之间的关系

样本土壤中的含水量范围为 5.0%–10.6%, pH 值为 7.5–9.0, 土壤含氮盐 NO_3^- 和 NH_4^+ 的浓度变化范围分别为 0.51–8.55 $\mu\text{g/g}$ 干重土壤和 3.60–6.51 $\mu\text{g/g}$ 干重土壤, 而 SO_4^{2-} 和 Cl^- 离子浓度变化范围较大, 分别在 19.6–308.5 $\mu\text{g/g}$ 干重土壤和 5.9–52.5 $\mu\text{g/g}$ 干重土壤。同时根据 16S rRNA 基因测序结果获得总微生物的组成信息, 使用 SPSS 软件进行 Pearson 及 Spearman 两种相关性分析方法分析环境因子与微生物分布之间的相关性, 分析结果如表 1 所示。从表 1 中可知, 含水量、pH、氮盐、 SO_4^{2-} 及 Cl^- 均未表现出与微生物分布的相关性。

2.2 土壤微生物多样性分析

从杭锦旗地区 32 个样本中共鉴定出 5 379 046 条高质量的 16S rRNA 基因序列, 高质量序列读长主要集中在 390–430 bp。经嵌合

体检测和去除后, 共获得 4 748 569 个有效标签。大多数序列(99.73%)可归类为细菌门。通过 97.00%的序列相似性定义了 2 294 个操作性分类单元(OTU)。其中包括丰度大于 4.00%的 6 个门: 放线菌门(*Actinobacteria*, 54.26%)、变形菌门(*Proteobacteria*, 18.21%)、绿屈挠菌门(*Chloroflexi*, 7.19%)、浮霉状菌门(*Planctomycetes*, 4.64%) 和 酸 杆 菌 门(*Acidobacteria*, 4.61%), 共占总序列的 88.91%。比较新召、什股壕气区和背景区样品门类丰度, 发现放线菌门在气区样品中更加丰富, 而变形菌门在背景区样品中相对丰度较高(图 2)。在对来自石油污染土壤中不同深度(20–120 cm)的 18 个分离株确定在原油基培养基(*crude oil-based medium*, COBM)上的生长能力的研究中, 发现放线菌具有最高的生长能力, 这与放线菌门能够降解烃类物质密切相关^[15]。在油藏中建立了嗜热(55 °C)正链烷烃降解($\text{C}_{21}\text{--}\text{C}_{30}$)培养物的研究中, 经过 900 多天的培育, 发现细菌群落由一个放线菌类谱系“*Candidatus Syntrophaliphaticia*”主导^[16]。

在属水平上包括相对丰度大于 2.00%的 9 个属: 类诺卡氏菌属(*Nocardioides*, 16.46%)、红色杆形菌属(*Rubrobacter*, 8.70%)、土壤红色杆形菌属(*Solirubrobacter*, 7.87%)、芽孢杆菌属(*Bacillus*, 5.37%)、盖亚女神菌属(*Gaiella*, 4.04%)、链霉菌

表 1 理化参数与微生物分布的相关关系

Table 1 The relationship between physicochemical parameters and the distribution of microorganisms

Physicochemical parameters	Pearson correlation coefficient	Significance	Spearman correlation coefficient	Significance
Water-content	0.005	0.965	0.134	0.239
pH	0.061	0.590	-0.178	0.117
NO_3^-	-0.137	0.230	-0.173	0.128
NH_4^+	0.050	0.661	-0.104	0.364
SO_4^{2-}	-0.199	0.138	0.023	0.864
Cl^-	-0.205	0.126	-0.26	0.051

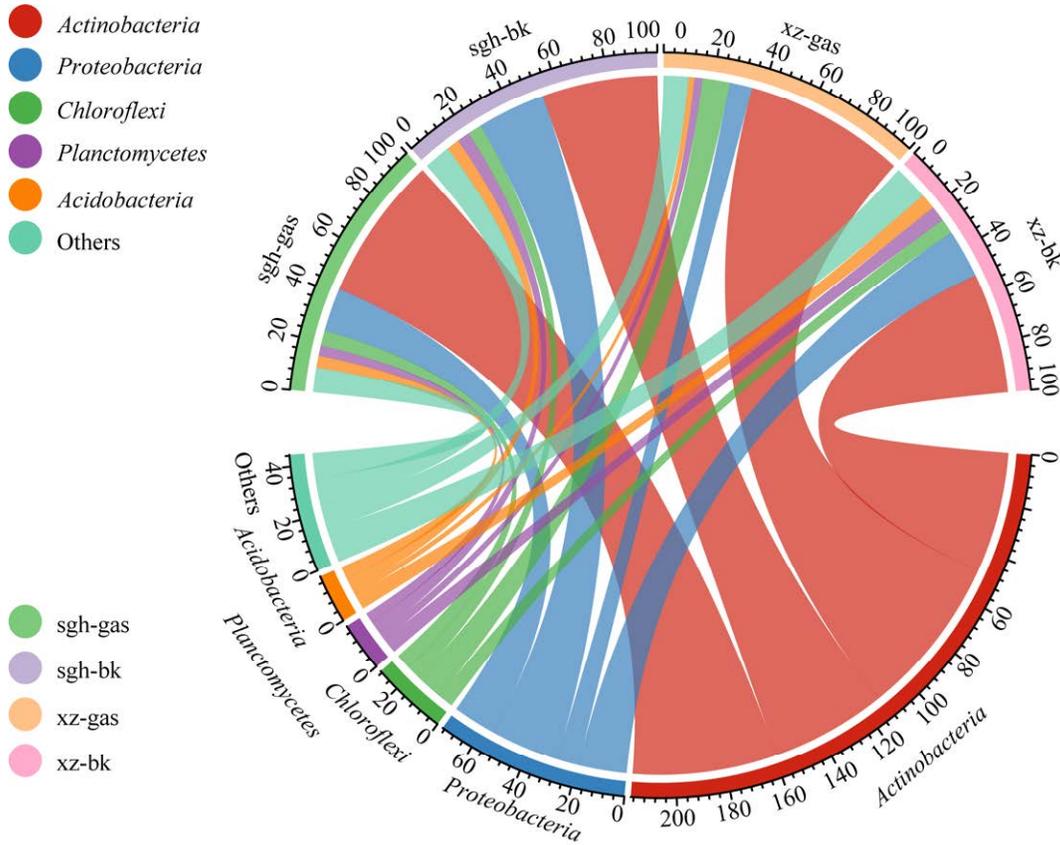


图 2 杭锦旗地区微生物在门水平上的分布

Figure 2 Taxonomic distribution of microbes at phylum level in Hangjinqi area.

属(*Streptomyces*, 3.40%)、气微菌属(*Aeromicrobium*, 2.85%)、土壤单胞菌属(*Solimonas*, 2.62%)和假诺卡氏菌属(*Pseudonocardia*, 2.21%)。选取丰度排名前 50 的属, 根据其在每个样品中的丰度信息, 在物种和样品两个层面进行聚类分析, 发现两个区域的菌群存在明显的异常富集现象(图 3)。在什股壕气区的样品中, 类诺卡氏菌属和气微菌属呈现显著富集现象。而在新召气区则为芽孢杆菌属。此外, 红色杆形菌属和土壤红色杆形菌属在这两个区域的样品中都显示出明显的富集特征。

基于 Shannon 指数和 Simpson 指数评估样本 α 多样性。Shannon 指数描述了样本中物种的丰度和均匀度, 新召背景区的 Shannon 指数最高,

并且与新召气区存在显著差异, 而什股壕气区和背景区之间的 Shannon 指数不存在差异(图 4A)。Simpson 指数反映了高丰度物种的分布情况, 发现不同组之间的 Simpson 指数存在类似的差异, 并且 4 组的辛普森指数都较大, 说明这 4 组样本中均存在许多高丰度物种(图 4B)。

群落的 β 多样性反映了不同样品中微生物群落组成的差异程度。本研究基于 unweighted-unifrac 距离对 4 种样本进行了 PCoA 分析, 如图 5 所示, 所有样本在水平方向上聚集为两个簇, 气区样本聚集在左侧, 而背景区样本聚集在右侧。水平坐标轴解释了样本间总体差异的 41.83%。研究表明, 杭锦旗地区气区和背景区的微生物群落多样性和组成具有明显差异。

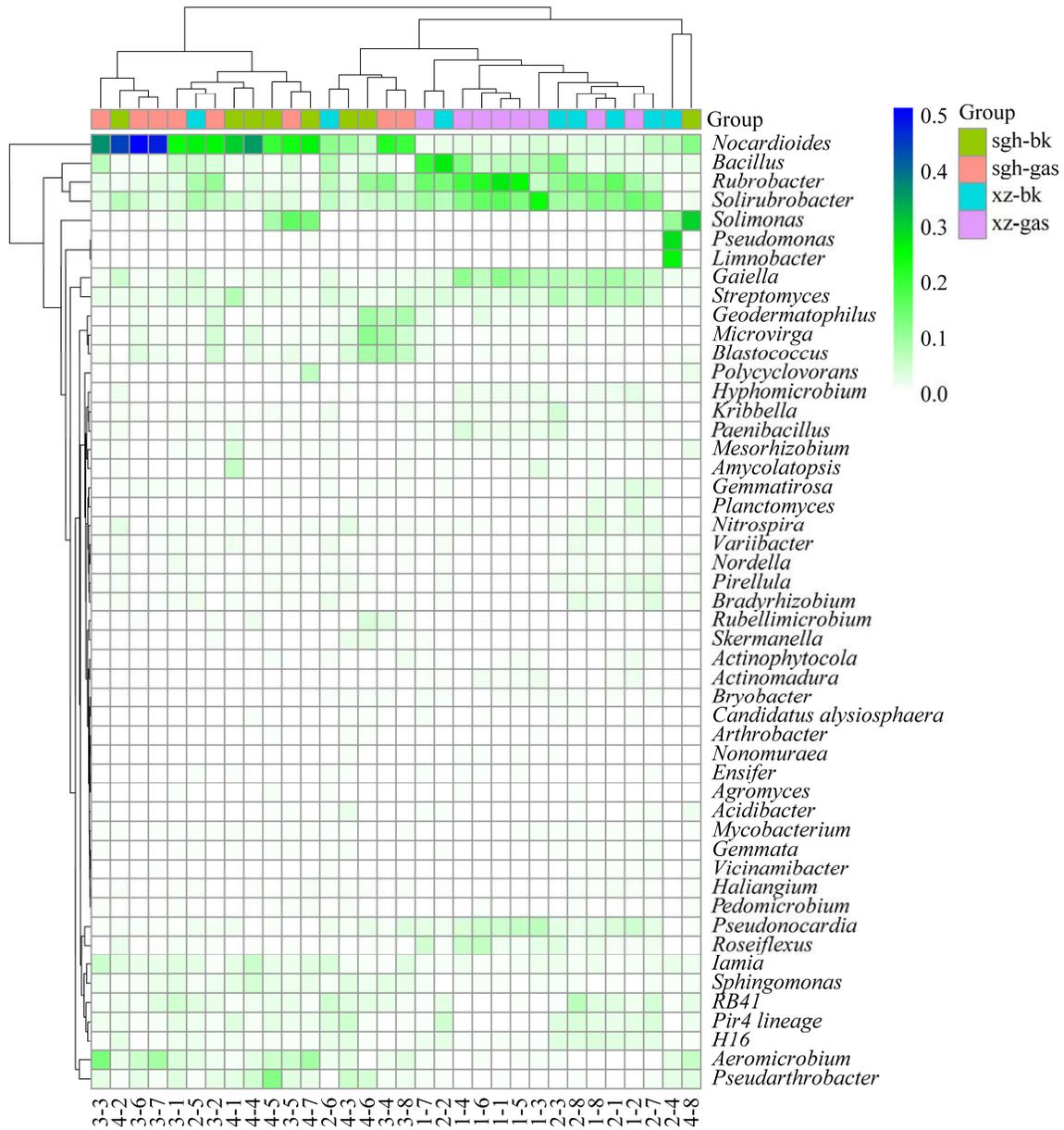


图 3 杭锦旗地区微生物属水平丰度聚热图
 Figure 3 Heat distribution at genus level of microbes in Hangjinqi area.

2.3 差异物种分析

为了比较不同区域的微生物群落组成,采用了曼哈顿图和 LEfSe 分析两种方法。LEfSe 分析不仅利用了非参数 Kruskal-Wallis 秩和检测以及 Wilcoxon 秩和检验,寻找不同分组间丰度差异显著的物种,而且还采用了线性判别分析(linear discriminant analysis, LDA)来估算每个组分(物

种)丰度对差异效果的影响大小。相比于曼哈顿图分析,LEfSe 分析方法具有更高的分辨率和准确性,可以更精确地分析不同区域微生物群落的差异。根据样本中门水平到属水平的微生物丰度进行 LEfSe 分析, LDA 得分较高的细菌类群表明在某一分组中的丰度较高,由于某些分组的差异微生物过多,选取较大的 LDA 值绘制图形。

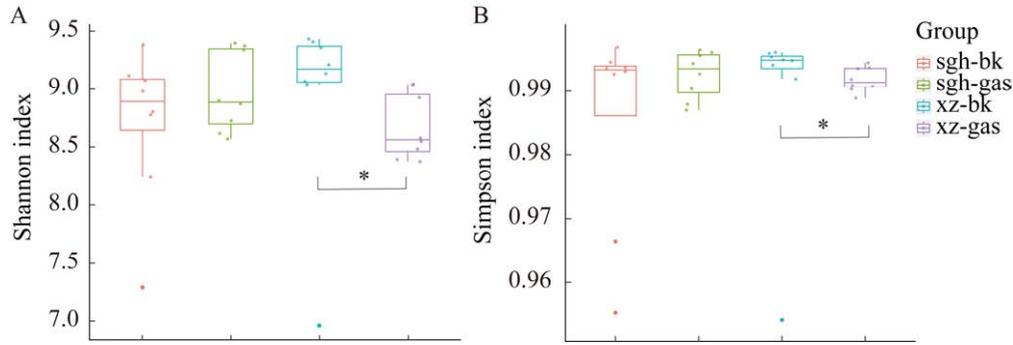


图 4 杭锦旗地区的 α 多样性指数

Figure 4 Alpha diversity index in Hangjinqi area. Assessing sample alpha diversity based on Shannon index (A) and Simpson index (B). *: $P < 0.05$.

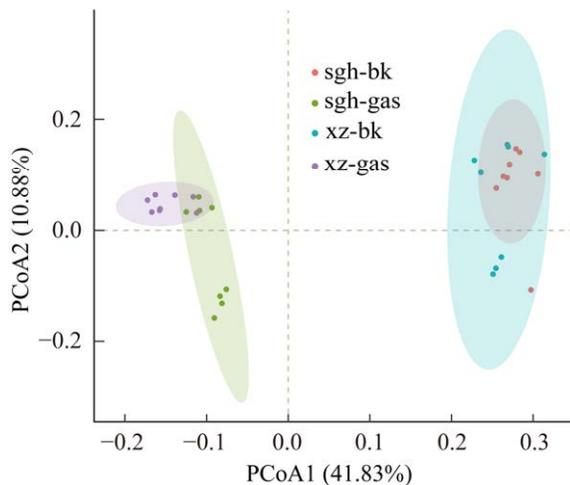


图 5 杭锦旗地区样本基于 unweighted-unifrac 距离的 PCoA 分析

Figure 5 PCoA analysis based on unweighted-unifrac distance of samples in Hangjinqi area.

在门水平上,曼哈顿分析结果显示新召气区显著富集的 OTUs 主要属于放线菌门和绿屈挠菌门(图 6A),而什股壕气区显著富集的 OTUs 数量明显减少,而且主要属于放线菌门和变形菌门(图 6B),LEfSe 分析结果显示新召气区中变形菌门和绿屈挠菌门的 LDA 值分别为 4.91 和 4.41 (图 7A、7B),什股壕气区相较于之前的分析,得到的差异菌属较少(图 7C、7D),其中只有变形菌门 LDA 值为 4.31。基于 OTUs 的曼哈顿分析结果与 LEfSe 分析结果基本一致。

在属水平上,曼哈顿分析结果显示新召气区显著富集 OTUs 主要包括土壤红色杆形菌属、球形孢囊菌属(*Sphaerisporangium*)、假诺卡氏菌属和生丝微菌属(*Hyphomicrobium*),而什股壕气区则主要包括类诺卡氏菌属、红色杆形菌属、盖亚女神菌属、小土壤杆菌属(*Parviterribacter*)和出芽单胞菌属(*Gemmatimonas*)。同时还对两个气区样本进行了曼哈顿分析,发现新召气区显著富集 OTUs 为红色杆形菌属、土壤红色杆形菌属、生丝微菌属(图 6C),而什股壕气区为类诺卡氏菌属、地嗜皮菌属(*Geodermatophilus*)、假节杆菌属(*Pseudarthrobacter*)。LEfSe 分析结果显示,新召气区 LDA 值大于 2 的菌属包括土壤红色杆形菌属、假诺卡氏菌属、生丝微菌属、马杜拉放线菌属(*Actinomadura*)、球形孢囊菌属、弗林德斯菌属(*Flindersiella*)和科恩氏菌属(*Cohnella*),其 LDA 值分别为 4.03、3.58、3.20、3.05、2.51、2.46 和 2.41,这与曼哈顿图分析得到的结果一致。然而什股壕气区 LDA 值大于 2 的菌属只有红色杆形菌属,其 LDA 值为 3.91。同时对两个气区样品进行了 LEfSe 分析,发现新召气区的显著富集菌属为红色杆形菌属、土壤红色杆形菌属,其 LDA 值为 4.17 和 4.09,而什股壕气区则为类诺卡氏菌属和气微菌属, LDA 值为 4.94 和 4.18 (图 7E、7F)。

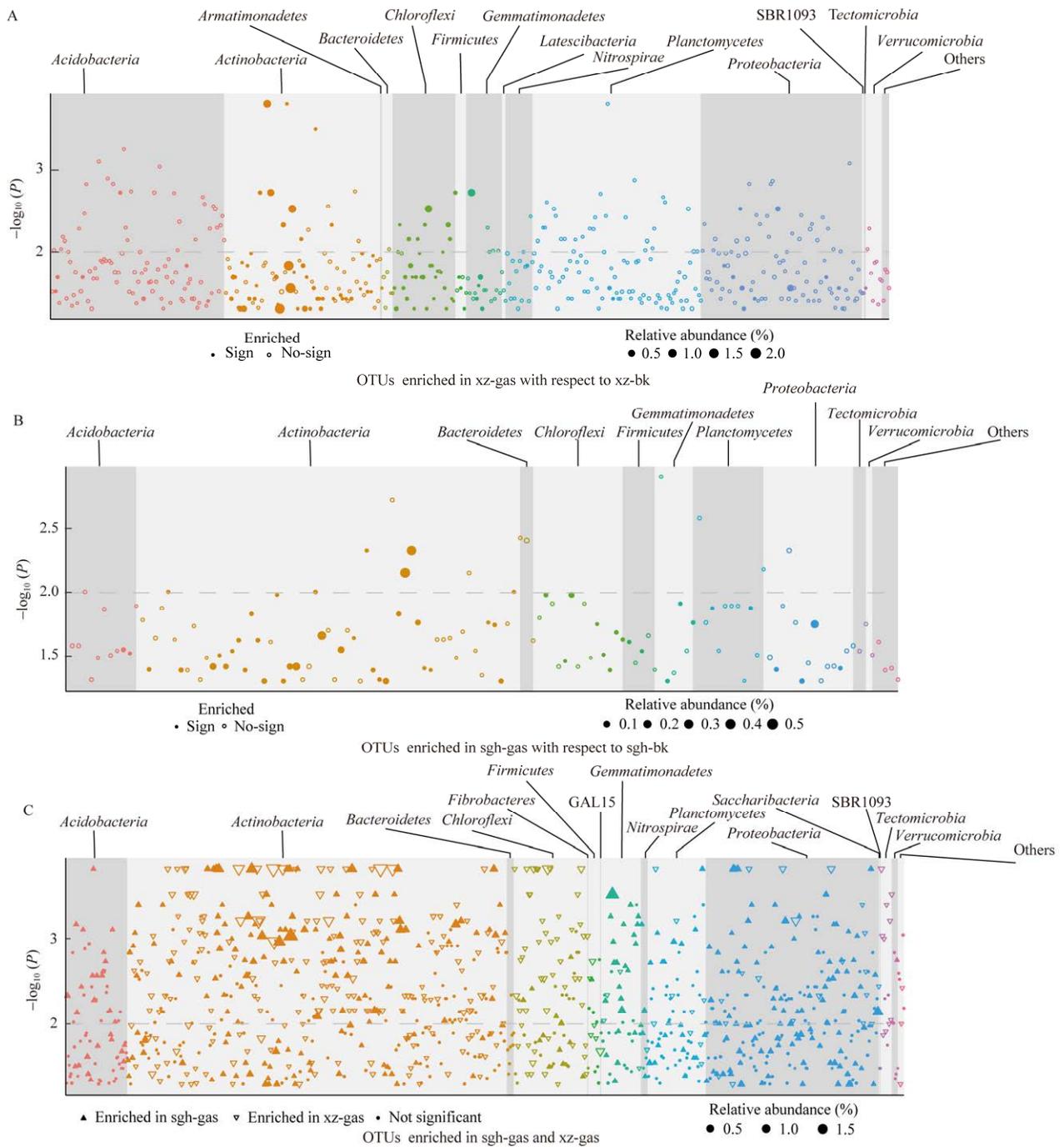


图 6 曼哈顿图显示显著富集 OTU

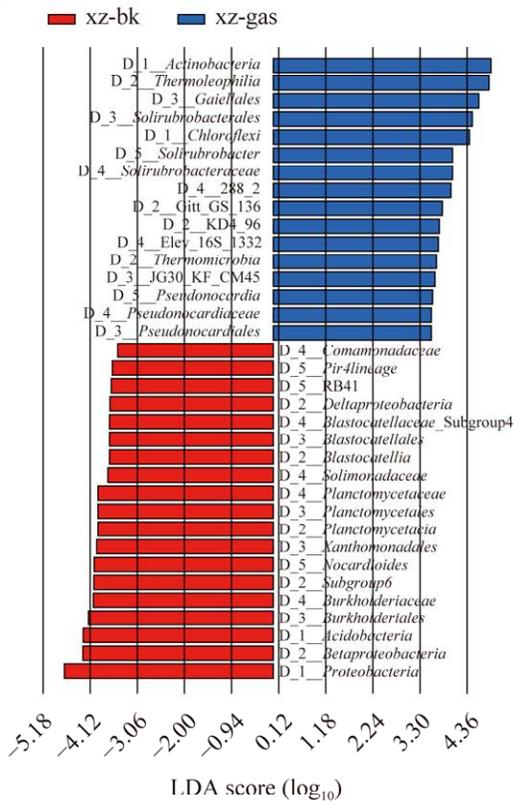
Figure 6 Manhattan plot displays significant enrichment of OTUs. A: OTUs enriched in xz-gas with respect to xz-bk. B: OTUs enriched in sgh-gas with respect to sgh-bk. C: OTUs enriched in sgh-gas and xz-gas.

2.4 微生物共生网络分析

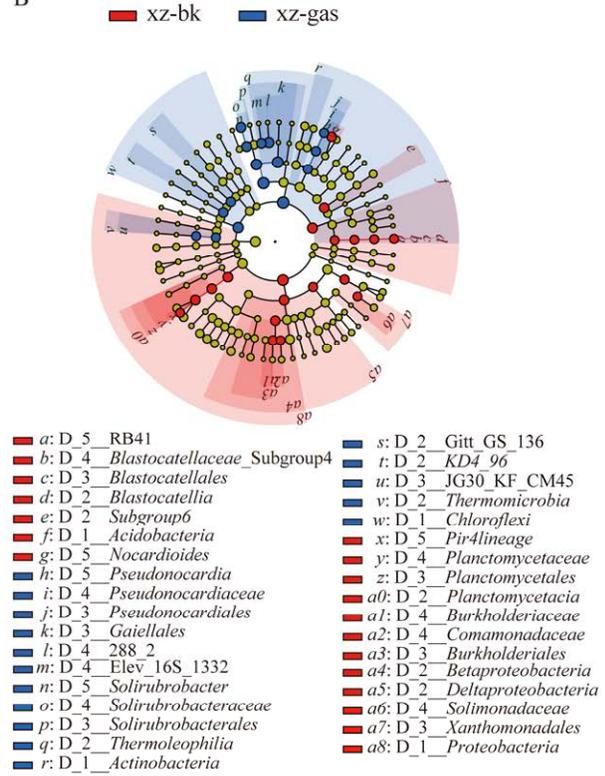
选取新召和什股壕气区样品进行微生物共生网络分析来探究微生物之间的相互关系。首先过滤掉相对丰度小于 0.5% 的微生物属，并进

行 Spearman 相关性分析，筛选出相关系数大于 0.7 且显著性 $P < 0.05$ 的相关关系构建相关性网络。生成的网络图包含 95 个节点(代表微生物属)和 460 条边,平均每个节点有 4.842 条边连接(图 8)。

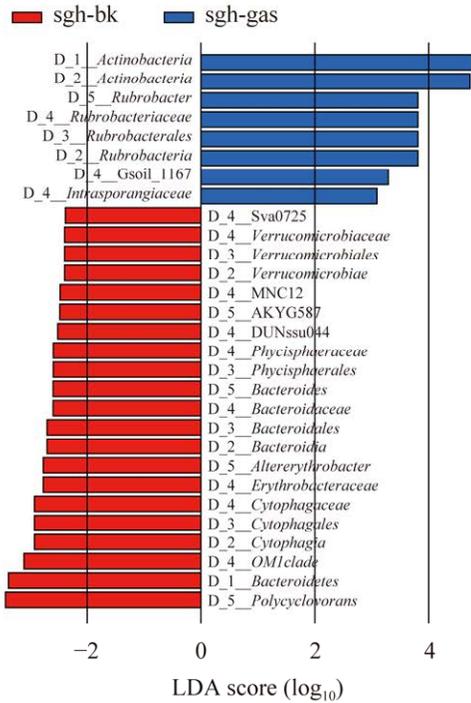
A



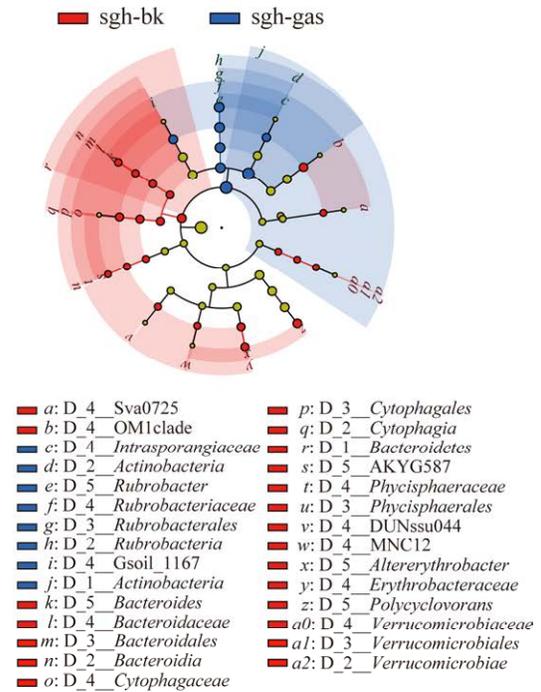
B



C



D



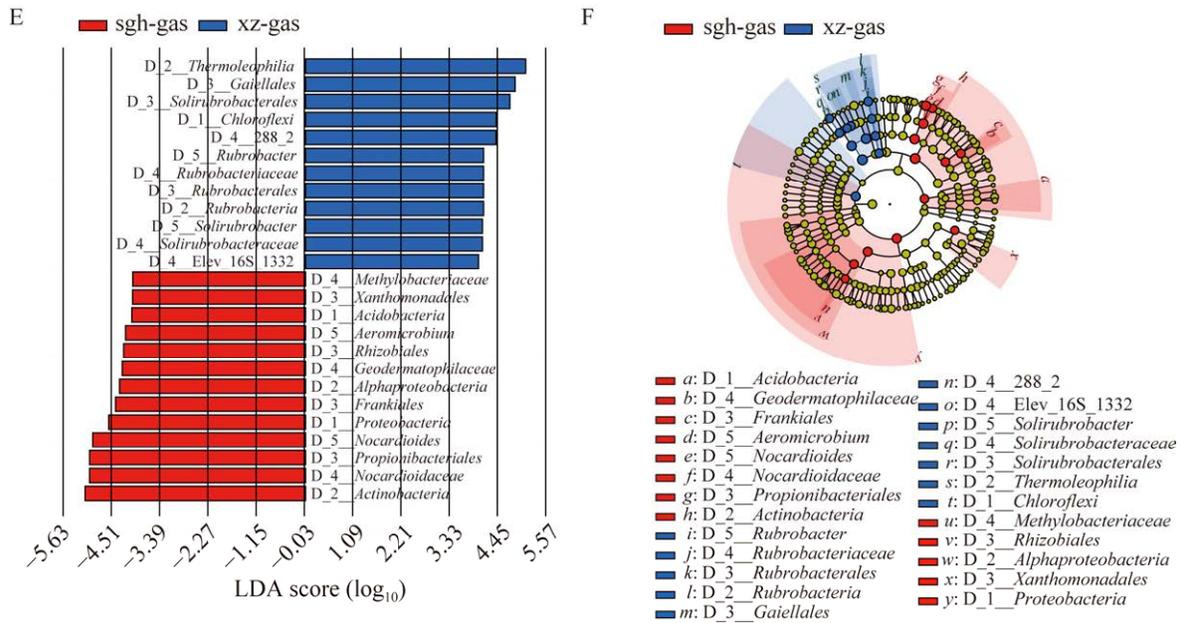
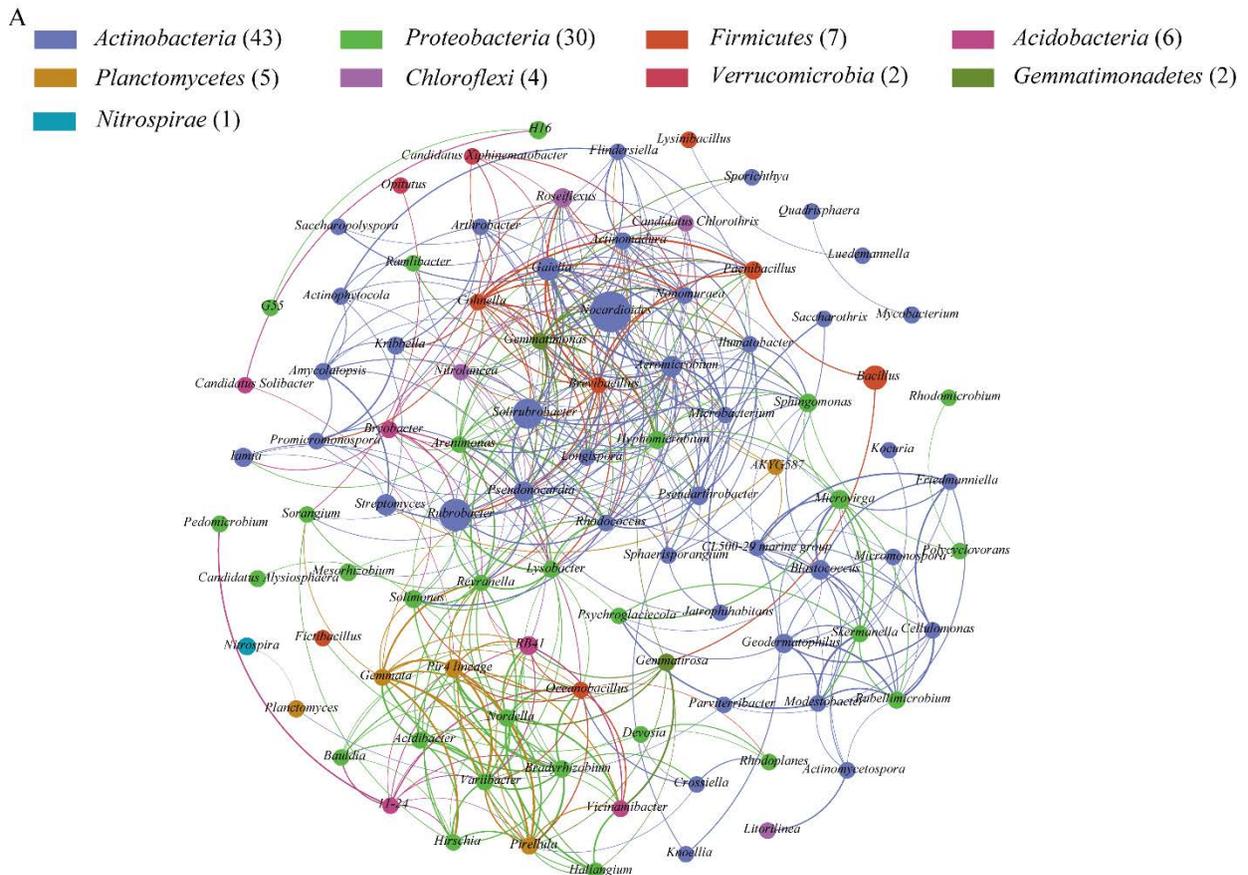


图 7 新召(A 和 B)、什股壕(C 和 D)、气区(E 和 F)微生物丰度的 LefSe 分析

Figure 7 LefSe analysis of microbial abundance between xz-gas and xz-bk (A and B), sgh-gas and sgh-bk (C and D), sgh-gas and xz-gas (E and F). A, C, E are LDA score with a threshold value of 3.5, 2.0, 4.0; B, D, F is the cladogram of microbial communities.



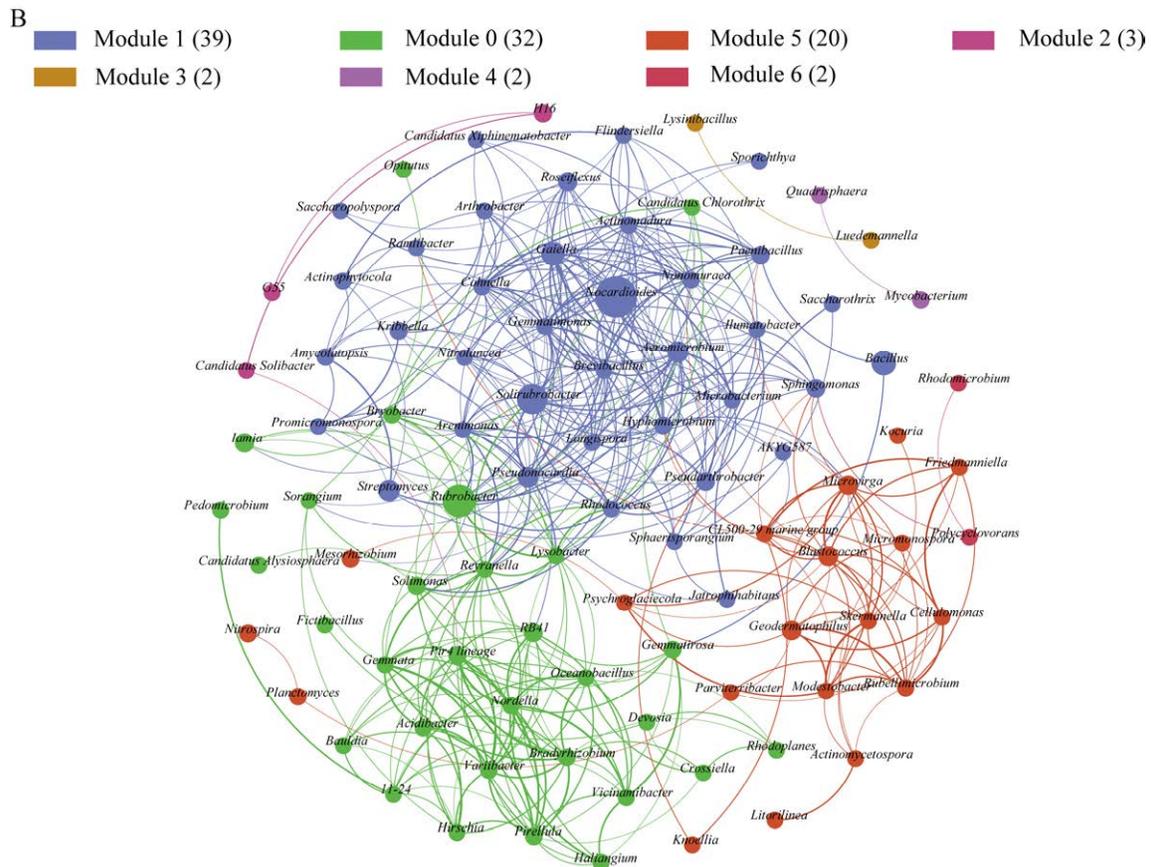


图 8 共发生网络分析 A: 根据门对节点着色. B: 根据模块对节点着色

Figure 8 Co-occurrence network analysis. Node size represents relative abundance, and edge thickness indicates the strength of correlation. A: Nodes colored based on phylum. B: Nodes colored based on modules.

除此之外，还计算了网络的拓扑特性参数，并与同等大小的 Erdős-Rényi 随机网络图进行了比较。拓扑特性参数包括平均路径长度、聚类系数和模块性指数。在的微生物网络中，平均路径长度为 3.171，比随机网络的平均路径长度 (2.228) 更长。这意味着微生物网络中节点之间的平均距离较大，信息传播的路径更长。聚类系数是衡量节点聚集程度的指标。在的网络中，聚类系数为 0.585，而随机网络的聚类系数仅为 0.109。这表明微生物网络中的节点更倾向于形成紧密的社区或群集。模块性指数是衡量网络中的模块化程度的指标。微生物网络具有模块性指数为 0.474，而随机网络的模块性指数

为 0.263。当模块性指数大于 0.4 时，表示网络具有明显的模块结构。此外，网络图中的节点主要分布在 5 个门，包括放线菌门、变形菌门、厚壁菌门 (*Firmicutes*)、酸杆菌门和浮霉状菌门。这些节点形成了 3 个主要的模块，这表明微生物之间可能存在特定的相互作用和功能关联。

3 讨论

本研究通过将微生物相对丰度数据降维成一维数据，最大程度地保留了物种相对丰度信息缩减相关分析的复杂程度，通过评估理化参数与微生物分布的相关性发现微生物的分布几

乎不受理化参数的干扰,避免了理化参数引起的微生物异常。同时根据多样性分析发现气区样品在门水平上呈现为放线菌门、变形菌门的富集,在属水平则为芽孢杆菌属、类诺卡氏菌属、鞘氨醇单胞菌属(*Sphingomonas*)和假单胞菌属(*Pseudomonas*)。先前的报道中发现这些菌属都具有降解石油碳氢化合物的能力^[17-21]。通过进一步的曼哈顿图分析和 LEfSe 分析,发现新召地区和什股壕地区均表现出了放线菌门的显著差异,这与之前放线菌门是最多功能和高效的可培养的碳氢化合物降解菌的报道相符合^[22]。本研究得到新召气区的指示菌属为土壤红色杆形菌属、球形孢囊菌属、假诺卡氏菌属和生丝微菌属。在一项利用 16S rRNA 基因测序分析石油污泥微生物群落的研究中检测出土壤红色杆形菌属是优势菌属^[23]。另外,在一项利用同位素探针(isotope probes, SIP)调查农业土壤中产甲烷条件下的甲苯降解菌的研究中,研究人员发现在重质 DNA 分离物中,假诺卡氏菌属和消化链球菌科(*Peptostreptococcaceae*)的菌群富集程度较低,他们推测这些菌群能够利用烃类中间产物,参与甲苯的降解过程^[24]。生丝微菌属在原油处理的土壤中是优势菌属,能够降解土壤环境中的烃类物质^[25]。球形孢囊菌属在之前的报道中未被发现与油气指示有关,可能得到一种新的油气指示菌属。什股壕气区的指示菌属则为类诺卡氏菌属、红色杆形菌属。类诺卡氏菌属尽管在 LEfSe 分析结果中未被发现,但该种菌属在什股壕气区显著富集,而且具有降解石油碳氢化合物的能力。此外, Deng 等在大港油田的石油烃污染土壤中检测到了红色杆形菌属等优势土著微生物的代谢活动,这些微生物对不同电解时间的生物处理垃圾填埋渗滤液中有机物组分的差异具有显著影响^[26]。在研究不同气区样品之间的显著差异物种时,新召气区

显著富集的为红色杆形菌属和土壤红色杆形菌属,而什股壕气区则为类诺卡氏菌属和气微菌属,这些菌属都是油气指示微生物,能够利用土壤中的石油碳氢化合物维系自身的生长发育。这表明了在不同地区的油气藏上方的微生物多样性存在显著差异。例如,在中国的 22 个不同油藏中,已发现细菌和古菌的多样性显著不同,这种多样性差异主要受到勘探模式、地层的温度、盐度和 pH 值等因素的影响^[27]。

在受碳氢化合物影响的地点,微生物展示的关键生化过程包括硫酸盐还原、发酵、甲烷生成、NO₃⁻还原以及铁和锰还原^[28]。因此,希望通过气区样品之间的共发生网络来探究杭锦旗地区微生物之间的相互作用和功能关联。本研究发现放线菌门和变形菌门是主导门,共同形成了三大主要模块。选取其中特征向量中心性排名前六的菌属:出芽单胞菌属、土壤红色杆形菌属、假诺卡氏菌属、短芽孢杆菌属(*Brevibacillus*)、气微菌属和类诺卡氏菌属。已有报道指出土壤中的石油碳氢化合物降解与土壤红色杆形菌属、假诺卡氏菌属以及类诺卡氏菌属相关^[18,23,24]。出芽单胞菌属在之前的报道中被发现具有还原 N₂O 的代谢潜力,能够影响土壤中的氮循环路径^[29]。短芽孢杆菌属能够产生脂肽类化合物,并具有增强油田采收效果的能力^[30]。气微菌属在原油泄漏饱和土壤的生物修复的某个阶段是优势菌种^[31]。因此,可以推测上述 6 类菌种在杭锦旗气区土壤中的烃类物质降解过程中扮演着重要角色。同时,将共发生网络与随机网络比较,发现气区样品的细菌群落具有非随机性与连通性,这表明确定性在微生物群落的构建过程中起主导作用。此外,可由模块化的共发生网络中发现,该微生物群落形成三大主要模块,不同的模块中的节点可能执行不同的功能。模块 1 中可能参与土壤中碳和氮的循环过程。如硝酸盐还原菌属

(*Nitrolancea*)为硝化螺菌纲(*Nitrospira*), 通过将亚硝酸盐(NO_2^-)氧化为硝酸盐(NO_3^-), 在氮循环中起重要作用^[32]。类诺卡氏菌属可以参与硝化、反硝化和固氮等过程, 有助于生态系统中氮的循环和可利用性, 同时还参与有机物质的分解, 将碳释放回环境中^[33]。模块 0 可能参与土壤中有有机物质的降解, 如红色杆形菌属在废水中有机物质的降解中起重要作用^[34]。土壤单胞菌属菌株中含有加氧酶, 能够在细胞外将聚合物裂解为低分子量产物, 这些产物可被吸收并用作碳源^[35]。

本研究通过 β 多样性分析发现气区和背景区样品在第一轴很好地区分开, 这表明油气区与背景区地表土壤的微生物群落结构差异能够通过大数据分析的方法被识别出来。同时, 利用差异分析得到的油气指示菌属在气区样品中并不是优势菌属, 相对丰度不超过 1%, 利用大数据分析的方法能够检测出低丰度油气指示微生物。因此, 大数据分析的方法可以很好地被用于油气微生物勘探, 能够表征油气区地表土壤的低丰度特异性指示菌属。

4 结论

本研究通过高通量测序的方法, 对杭锦旗气区和背景区微生物群落组成和结构进行分析, 通过相关分析排除了理化参数对微生物分布的干扰, 结合共生网络分析, 发现杭锦旗气区和背景区的微生物分布具有显著差异, 而且不同气区的指示微生物存在差异, 新召气区的指示菌属为土壤红色杆形菌属、球形孢囊菌属、假诺卡氏菌属和生丝微菌属, 而什股壕气区为类诺卡氏菌属、红色杆形菌属。通过共发生网络分析发现, 气区的微生物群落构建过程是非随机的, 而且存在主要功能模块, 通过对模块的特征性菌属分析, 认为存在碳氮循环模块和有机物质降解模块。本研究通过大数据分析的方法对典型气区的指示微

生物进行分析, 能够很好地表征气区的微生物群落结构特征, 可以用于检测一些低丰度的油气指示菌属, 具有良好的油气勘探应用前景。

参考文献

- [1] SAUNDERS DF, RAY BURSON K, KEITH THOMPSON C. Model for hydrocarbon microseepage and related near-surface alterations[J]. AAPG Bulletin, 1999, 83(1): 170-185.
- [2] RASHEED MA, LAKSHMI M, SRINU D, DAYAL AM. Bacteria as indicators for finding oil and gas reservoirs: a case study of the Bikaner-Nagaur Basin, Rajasthan, India[J]. Petroleum Science, 2011, 8(3): 264-268.
- [3] RASHEED MA, HASAN SZ, SRINIVASA RAO PL, BORUAH A, SUDARSHAN V, KUMAR B, HARINARAYANA T. Application of geo-microbial prospecting method for finding oil and gas reservoirs[J]. Frontiers of Earth Science, 2015, 9(1): 40-50.
- [4] ZHANG CY, HE Z, ZHANG S, YIN MY, NING Z, LIU YC. A DNA-based analysis of a microbial technique for the prospecting of oil and gas applied to a known oil field, China[J]. Geomicrobiology Journal, 2017, 34(1): 63-70.
- [5] RASHEED MA, PRASANNA MV, KUMAR T, PATIL DJ, DAYAL AM. Geo-microbial prospecting method for hydrocarbon exploration in Vengannapalli Village, Cuddapah Basin, India[J]. Current Science, 2008, 95: 361-366.
- [6] RASHEED MA, KALPANA MS, VEENA PRASANNA M, LAKSHMI M, MADHAVI T, MANI TIWARI D, PATIL DJ, DAYAL AM, RAJU SV. Geo-microbial and light gaseous hydrocarbon anomalies in the near surface soils of Deccan Syncline Basin, India: implications to hydrocarbon resource potential[J]. Journal of Petroleum Science and Engineering, 2012, 84/85: 33-41.
- [7] RASHEED MA, LAKSHMI M, KALPANA MS, PATIL DJ, DAYAL AM. Recognition of hydrocarbon microseepage using microbial and adsorbed soil gas indicators in the petroliferous region of Krishna-Godavari Basin, India[J]. Current Science, 2017, 112(3): 560-568.
- [8] 汤玉平, 许科伟, 顾磊, 杨帆, 高俊阳, 任春, 王国建. 油气微生物勘探理论与技术研究进展[J]. 石油实验地质, 2021, 43(2): 325-334.

- TANG YP, XU KW, GU L, YANG F, GAO JY, REN C, WANG GJ. Recent progress in the theory and technology of microbial prospecting for oil and gas[J]. *Petroleum Geology and Experiment*, 2021, 43(2): 325-334 (in Chinese).
- [9] SHENDURE J, BALASUBRAMANIAN S, CHURCH GM, GILBERT W, ROGERS J, SCHLOSS JA, WATERSTON RH. DNA sequencing at 40: past, present and future[J]. *Nature*, 2017, 550(7676): 345-353.
- [10] 褚厚娟, 郝纯, 邓诗财, 麻婷婷, 梅海. 木里地区水合物轻烃微渗漏微生物群落及烃氧化菌响应特征研究[J]. *物探与化探*, 2017, 41(6): 1037-1043.
CHU HJ, HAO C, DENG SC, MA TT, MEI H. A study of gas hydrate light hydrocarbon microseepage microbial communities and response characteristics of hydrocarbon oxidizing microorganism in Muli area[J]. *Geophysical and Geochemical Exploration*, 2017, 41(6): 1037-1043 (in Chinese).
- [11] 倪春华, 朱建辉, 刘光祥, 王付斌, 贾会冲, 张威, 武英利, 缪九军. 鄂尔多斯盆地杭锦旗地区上古生界煤系烃源岩生烃潜力再评价[J]. *石油实验地质*, 2021, 43(5): 826-834.
NI CH, ZHU JH, LIU GX, WANG FB, JIA HC, ZHANG W, WU YL, MIAO JJ. Re-evaluation of hydrocarbon generation potential of the Upper Paleozoic coal-measure source rocks in the Hangjinqi area of Ordos Basin[J]. *Petroleum Geology and Experiment*, 2021, 43(5): 826-834 (in Chinese).
- [12] 孙乃泉, 李良, 蒲仁海, 彭胜琴. 鄂尔多斯盆地北部杭锦旗地区中元古界断陷—拗陷层地层特征[J]. *物探与化探*, 2020, 44(5): 1144-1152.
SUN NQ, LI L, PU RH, PENG SQ. The fault and depression characteristics of Mesoproterozoic strata in Haggin Banner area, northern of Ordos Basin[J]. *Geophysical and Geochemical Exploration*, 2020, 44(5): 1144-1152 (in Chinese).
- [13] 许科伟, 郑旭莹, 顾磊, 郭嘉琪, 汤玉平. 油气微生物分子勘探技术与初步应用[J]. *石油实验地质*, 2022, 44(5): 896-903.
XU KW, ZHENG XY, GU L, GUO JQ, TANG YP. Molecular microbial prospecting for oil and gas and its preliminary application[J]. *Petroleum Geology and Experiment*, 2022, 44(5): 896-903 (in Chinese).
- [14] CAPORASO JG, LAUBER CL, WALTERS WA, BERG-LYONS D, LOZUPONE CA, TURNBAUGH PJ, FIERER N, KNIGHT R. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2011, 108(supplement_1): 4516-4522.
- [15] OLAJUYIGBE FM, EHIOSUN KI. Assessment of crude oil degradation efficiency of newly isolated actinobacteria reveals untapped bioremediation potentials[J]. *Bioremediation Journal*, 2016, 20(2): 133-143.
- [16] LIU YF, CHEN J, LIU ZL, SHOU LB, LIN DD, ZHOU L, YANG SZ, LIU JF, LI W, GU JD, MU BZ. Anaerobic degradation of paraffins by thermophilic actinobacteria under methanogenic conditions[J]. *Environmental Science & Technology*, 2020, 54(17): 10610-10620.
- [17] SINGER AC, WONG CS, CROWLEY DE. Differential enantioselective transformation of atropisomeric polychlorinated biphenyls by multiple bacterial strains with different inducing compounds[J]. *Applied and Environmental Microbiology*, 2002, 68(11): 5756-5759.
- [18] HABE H, OMORI T. Genetics of polycyclic aromatic hydrocarbon metabolism in diverse aerobic bacteria[J]. *Bioscience, Biotechnology, and Biochemistry*, 2003, 67(2): 225-243.
- [19] BAEK KH, KIM HS, OH HM, YOON BD, KIM J, LEE IS. Effects of crude oil, oil components, and bioremediation on plant growth[J]. *Journal of Environmental Science and Health (Part A)*, 2004, 39(9): 2465-2472.
- [20] CHO O, CHOI KY, ZYLSTRA GJ, KIM YS, KIM SK, LEE JH, SOHN HY, KWON GS, KIM YM, KIM E. Catabolic role of a three-component salicylate oxygenase from *Sphingomonas yanoikuyae* B1 in polycyclic aromatic hydrocarbon degradation[J]. *Biochemical and Biophysical Research Communications*, 2005, 327(3): 656-662.
- [21] BELL TH, YERGEAU E, MARTINEAU C, JUCK D, WHYTE LG, GREER CW. Identification of nitrogen-incorporating bacteria in petroleum-contaminated Arctic soils by using [15N] DNA-based stable isotope probing and pyrosequencing[J]. *Applied and Environmental Microbiology*, 2011, 77(12): 4163-4171.
- [22] LIU JJ, WU J, LIN JW, ZHAO J, XU TY, YANG QC, ZHAO J, ZHAO ZM, SONG XF. Changes in the microbial community diversity of oil exploitation[J]. *Genes*, 2019, 10(8): 556.
- [23] ZHANG DC, MORTELMAIER C, MARGSIN R. Characterization of the bacterial archaeal diversity in hydrocarbon-contaminated soil[J]. *Science of the Total Environment*, 2012, 421/422: 184-196.

- [24] SUN WM, SUN XX, CUPPLES AM. Identification of *Desulfohalobos* as toluene-assimilating microorganisms from a methanogenic consortium[J]. *International Biodeterioration & Biodegradation*, 2014, 88: 13-19.
- [25] AL-JAWASIM M. Long-term combined effects of crude oil and dispersant on sediment bacterial community[J]. *Journal of Environmental Treatment Techniques*, 2020, 9(1): 259-263.
- [26] DENG Y, WEN JP, ZHU X, CHEN N, FENG CP, ZHENG YH, WANG HS, CHEN FX, GAO Y. Research on the redox behavior changes of humic-like substances wastewater during electrochemical oxidation process and using the treated effluent to improve the heavily contaminated soil: taking petroleum hydrocarbon contaminated soil as example[J]. *Journal of Cleaner Production*, 2020, 263: 121398.
- [27] GAO PK, TIAN HM, WANG YS, LI YS, LI Y, XIE JX, ZENG B, ZHOU JF, LI GQ, MA T. Spatial isolation and environmental factors drive distinct bacterial and archaeal communities in different types of petroleum reservoirs in China[J]. *Scientific Reports*, 2016, 6: 20174.
- [28] ZHAO F, LI P, GUO C, SHI RJ, ZHANG Y. Bioaugmentation of oil reservoir indigenous *Pseudomonas aeruginosa* to enhance oil recovery through *in situ* biosurfactant production without air injection[J]. *Bioresource Technology*, 2018, 251: 295-302.
- [29] CHEE-SANFORD J, TIAN D, SANFORD R. Consumption of N₂O and other N-cycle intermediates by *Gemmatimonas aurantiaca* strain T-27[J]. *Microbiology* (Reading, England), 2019, 165(12): 1345-1354.
- [30] DONG H, ZHENG AY, HE YL, WANG XT, LI Y, YU GM, GU YG, BANAT IM, SUN SS, SHE YH, ZHANG F. Optimization and characterization of biosurfactant produced by indigenous *Brevibacillus borstelensis* isolated from a low permeability reservoir for application in MEOR[J]. *RSC Advances*, 2022, 12(4): 2036-2047.
- [31] ALI N, DASHTI N, KHANAFER M, AL-AWADHI H, RADWAN S. Bioremediation of soils saturated with spilled crude oil[J]. *Scientific Reports*, 2020, 10: 1116.
- [32] JIANG H, YANG P, WANG Z, REN S, QIU JG, LIANG HR, PENG YZ, LI XY, ZHANG Q. Novel insights into overcoming nitrite oxidation bacteria acclimatization problem in treatment of high-ammonia wastewater through partial nitrification[J]. *Bioresource Technology*, 2021, 336: 125254.
- [33] ANDERSON CR, CONDRON LM, CLOUGH TJ, FIERS M, STEWART A, HILL RA, SHERLOCK RR. Biochar induced soil microbial community change: implications for biogeochemical cycling of carbon, nitrogen and phosphorus[J]. *Pedobiologia*, 2011, 54(5/6): 309-320.
- [34] PATEL D, BAPODRA SL, MADAMWAR D, DESAI C. Electroactive bacterial community augmentation enhances the performance of a pilot scale constructed wetland microbial fuel cell for treatment of textile dye wastewater[J]. *Bioresource Technology*, 2021, 332: 125088.
- [35] BIRKE J, JENDROSSEK D. *Solimonas fluminis* has an active latex-clearing protein[J]. *Applied Microbiology and Biotechnology*, 2019, 103(19): 8229-8239.