

# 呼和浩特市周边地下水 CPR 和 DPANN 类群的检测及多样性分析

姜凯<sup>1,2</sup>, 曹春玲<sup>3</sup>, 红雨<sup>\*1,2</sup>, 宋枚汐<sup>1</sup>, 杨鸿臻<sup>1</sup>, 陈乌日娜<sup>1</sup>, 刘昱昕<sup>1</sup>

1 内蒙古师范大学生命科学与技术学院, 内蒙古 呼和浩特 010022

2 内蒙古自治区高等学校蒙古高原生物多样性保护与可持续利用重点实验室, 内蒙古 呼和浩特 010022

3 内蒙古自治区农牧业技术推广中心, 内蒙古 呼和浩特 010010

姜凯, 曹春玲, 红雨, 宋枚汐, 杨鸿臻, 陈乌日娜, 刘昱昕. 呼和浩特市周边地下水 CPR 和 DPANN 类群的检测及多样性分析[J]. 微生物学通报, 2023, 50(12): 5320-5336.

JIANG Kai, CAO Chunling, HONG Yu, SONG Meixi, YANG Hongzhen, CHEN Wurina, LIU Yuxin. Diversity of CPR and DPANN in groundwater around Hohhot City[J]. Microbiology China, 2023, 50(12): 5320-5336.

**摘要:** 【背景】候选门级辐射类群(candidate phyla radiation, CPR)和 DPANN 是与绝大多数已知细菌和古菌具有显著差异的独特类群, 因发现较晚, 人们对其认识还非常有限。已知二者类群庞大、存在广泛, 但在不同生境中的多样性研究还较少, 生态功能尚属未知。【目的】分析不同地下水中 CPR 和 DPANN 的多样性, 探讨不同方法对 CPR 和 DPANN 检出及富集的影响。【方法】采用 0.1 μm 滤膜收集菌体和宏基因组测序的方法分析呼和浩特市周边 4 个不同地下水中 CPR 和 DPANN 的多样性。比较宏基因组测序与 16S rRNA 基因扩增子测序对 CPR 和 DPANN 检出的影响, 以及不同过滤方式和不同滤膜组合对 CPR 和 DPANN 富集的作用。【结果】4 个地下水样品 CPR 类群检出 33–64 个菌门, 相对丰度在 0.17%–1.67% 之间; DPANN 检出 1–7 个菌门, 相对丰度在 0.000 93%–0.071 00% 之间。对于 CPR 和 DPANN, 16S rRNA 基因 V3–V4 区扩增子测序仅检出了纳古菌门(Nanoarchaeota)。1.2 μm 与 0.1 μm 滤膜组合具有最好的 CPR 和 DPANN 富集效果, 在及时更换滤膜的过滤方式下相对丰度分别提高到 13.33% 和 0.58%。【结论】地下水中存在种类丰富但相对丰度较低的 CPR 和 DPANN 物种资源, 且不同地下水中的 CPR 和 DPANN 存在一定的差异。16S rRNA 基因 V3–V4 区扩增子测序会遗漏地下水中 CPR 和 DPANN 物种信息。在及时更换滤膜的过滤方式下采用不同孔径滤膜组合过滤会显著提高 CPR 和 DPANN 的相对丰度, 达到富集效果。本研究为后续 CPR 和 DPANN 物种资源、基因资源和天然产物资源的

资助项目: 内蒙古自治区自然科学基金(2020BS03006); 内蒙古师范大学基本科研业务费项目(2022JBQN091, 2022JBTD010); 内蒙古自治区科技计划(2020GG0039)

This work was supported by the Natural Science Foundation of Inner Mongolia Autonomous Region (2020BS03006), the Fundamental Research Funds for Inner Mongolia Normal University (2022JBQN091, 2022JBTD010), and the Science and Technology Program of Inner Mongolia Autonomous Region (2020GG0039).

\*Corresponding author. E-mail: hongyu@imnu.edu.cn

Received: 2023-05-15; Accepted: 2023-06-26; Published online: 2023-07-28

挖掘, 以及该类菌株的可培养工作奠定了基础。

关键词: 候选门级辐射类群(CPR); DPANN; 宏基因组测序; 16S rRNA 基因扩增子测序; 过滤富集

## Diversity of CPR and DPANN in groundwater around Hohhot City

JIANG Kai<sup>1,2</sup>, CAO Chunling<sup>3</sup>, HONG Yu<sup>\*1,2</sup>, SONG Meixi<sup>1</sup>, YANG Hongzhen<sup>1</sup>,  
CHEN Wurina<sup>1</sup>, LIU Yuxin<sup>1</sup>

1 College of Life Science and Technology, Inner Mongolia Normal University, Hohhot 010022, Inner Mongolia, China

2 Key Laboratory of Biodiversity Conservation and Sustainable Utilization in Mongolian Plateau for College and University of Inner Mongolia Autonomous Region, Hohhot 010022, Inner Mongolia, China

3 Agriculture and Animal Husbandry Technology Popularization Center of Inner Mongolia Autonomous Region, Hohhot 010010, Inner Mongolia, China

**Abstract: [Background]** Candidate phyla radiation (CPR) and DPANN are unique groups that are significantly different from most known bacteria and archaea, and the knowledge about them is limited due to the late discovery. The two groups are known to be large and widespread, while their diversity and ecological roles in different habitats remain to be studied. **[Objective]** To analyze the diversity of CPR and DPANN in different groundwater samples and the influences of different methods on the detection and enrichment of CPR and DPANN. **[Methods]** Metagenomic sequencing was employed to determine the diversity of CPR and DPANN in four different groundwater samples around Hohhot after filtration through a 0.1  $\mu\text{m}$  membrane. Metagenomic sequencing was compared with 16S rRNA gene amplicon sequencing regarding the detection of CPR and DPANN, and the effects of different filtration methods and membrane combinations on the enrichment of CPR and DPANN were compared. **[Results]** From the 4 groundwater samples, 33–64 phyla of CPR with the relative abundance of 0.17%–1.67% and 1–7 phyla of DPANN with the relative abundance of 0.000 93%–0.071% were detected. For CPR and DPANN, only *Nanoarchaeota* was detected by 16S rRNA gene V3–V4 amplicon sequencing. The combination of 1.2  $\mu\text{m}$  and 0.1  $\mu\text{m}$  filters showed the best enrichment effect on CPR and DPANN, and the relative abundance increased to 13.33% and 0.58%, respectively, by timely replacement of filters. **[Conclusion]** There were abundant resources of CPR and DPANN with low relative abundance in groundwater, and the distribution of CPR and DPANN varied in different groundwater samples. The 16S rRNA gene V3–V4 amplicon sequencing missed the information of CPR and DPANN in groundwater. The enrichment effect on CPR and DPANN can be significantly improved by the combination of filters with different pore sizes and the timely replacement of filters. The findings underpin the further exploration of species, gene, and natural product resources and the strain culture of CPR and DPANN.

**Keywords:** candidate phyla radiation (CPR); DPANN; metagenomic sequencing; 16S rRNA gene amplicon sequencing; enrichment by filtration

候选门级辐射类群 (candidate phyla radiation, CPR) 细菌和 DPANN 超门古菌 (由最初的 5 个菌门 *Candidatus* Diapherotrites、*Ca.* Parvarchaeota、*Ca.* Aenigmarchaeota、纳古菌门 (*Nanoarchaeota*) 和 *Ca.* Nanohaloarchaeota 的首字母命名而来) 是一类重要的微生物暗物质, 约占地球原核生物的一半<sup>[1]</sup>。二者的研究起步较晚, 分别于 2015 年<sup>[2]</sup>和 2013 年<sup>[3]</sup>首次定名。随着宏基因组测序技术的广泛应用, 已在地下水、饮用水、土壤、极端环境和动物体内等广泛生境中发现有 CPR 和 DPANN 的存在<sup>[1,4-8]</sup>, 二者类群目前已分别扩展至至少 75 个“*Candidatus*”细菌菌门和 12 个“*Candidatus*”古菌菌门<sup>[1,9-10]</sup>。CPR 和 DPANN 虽分属细菌域和古菌域, 但二者有很多共性特征, 包括细胞微小、基因组精简、缺少核心代谢、营共生生活和难培养等<sup>[1-3,9,11-12]</sup>。然而现阶段对 CPR 和 DPANN 的认识仍不足, 在生境及多样性、生态功能、进化及演变规律、培养方法和生活方式等方面有待系统而深入的研究。

CPR 和 DPANN 类群菌株检测和富集是后续深入研究的前提。两类群中的很多菌在 16S rRNA 基因中存在插入片段, 同时它们的 16S rRNA 基因高度可变且很多是在保守区, 易造成通用引物错配、不匹配等现象的发生<sup>[1,3,13-16]</sup>, 导致在基于 16S rRNA 基因扩增的高通量测序中未能检出。CPR 和 DPANN 菌株细胞个体微小, 很多具有 0.22  $\mu\text{m}$  滤膜滤过性的特点<sup>[8,17-19]</sup>。上述两点使得常规基于 0.22  $\mu\text{m}$  滤膜收集菌体、16S rRNA 基因扩增子测序的微生物多样性研究不能完全反映 CPR 和 DPANN 菌在不同环境特别是水生环境中的存在及多样性。随着宏基因组测序成本的降低, 该技术的广泛应用为 CPR 和 DPANN 的研究尤其是多样性研究提供了前提条件。地下水是相对稳定的寡营养环

境, 研究显示富含 CPR 和 DPANN 菌株, 然而与其他细菌和古菌相比丰度仍较低<sup>[1]</sup>。尽可能多地收集 CPR 和 DPANN 菌株、提高相对丰度、减少其他细菌和古菌的干扰是后续深度测序拼接 CPR 和 DPANN 新物种基因组、挖掘其中天然产物资源、提高两类群菌株分离成功率的关键。

本研究以不同地下水为研究对象, 采用宏基因组测序方法对其中的 CPR 和 DPANN 多样性进行分析, 比较宏基因组测序与 16S rRNA 基因扩增子测序对 CPR 和 DPANN 检出的影响, 同时采用不同滤膜组合及不同过滤方式处理地下水样品, 探索 CPR 和 DPANN 菌富集的方法, 为下一步对 CPR 和 DPANN 菌株基因及天然产物资源的挖掘、可培养及生态功能研究提供参考。

## 1 材料与方法

### 1.1 样品采集

2022 年 5 月至 10 月之间, 分别在呼和浩特市地理位置相距较远的 4 个点采集地下水。地下水 1 (#1, 40°14'15"N, 112°11'41"E) 位于和林格尔县新店子镇东沟村, 为村民自家深井井水, 采样时水温 6.8 °C, pH 6.5–7.0 (pH 试纸), 盐浓度 0.1%, 井水微黄, 采样时外界气温为 15 °C。地下水 2 (#2, 40°12'39"N, 112°00'58"E) 位于和林格尔县新店子镇店湾行政村农场自然村, 为小型牧场内井水, 采样时水温 9.6 °C, pH 6.5–7.0 (pH 试纸), 盐浓度 0.2%, 采样时外界气温为 23 °C。地下水 3 (#3, 40°54'04"N, 111°48'13"E) 位于呼和浩特市新城区讨思浩村, 为村集体灌溉井水, 采样时 pH 6.5–7.0 (pH 试纸), 采样时外界气温为 20 °C。地下水 4 (#4, 40°52'48"N, 111°38'43"E) 位于呼和浩特市市郊红山口村, 为村集体所有深井井水 (定时供

给村民生活用水), 采样时 pH 6.5–7.0 (pH 试纸), 盐浓度 0.5%, 采样时外界气温为 9 °C。

采集方法: 采样前将 1 L 试剂瓶于 121 °C 灭菌 20 min, 采样桶(15 L 和 25 L)用 75%乙醇杀菌。井水先放水 20 min 左右, 排尽管路中积存的水, 之后用样水清洗桶 3 次后采水样。样品采集后当天运回实验室, 避光放置, 2 d 内处理完成。

## 1.2 主要试剂和仪器

DNA 文库试剂盒 NEBNext<sup>®</sup> Ultra DNA Library Preokit for Illumina、NEB Next<sup>®</sup> Ultra<sup>™</sup> II FS DNA PCR-free Library Prep Kit 和 Phusion<sup>®</sup> High-Fidelity PCR Master Mix, New England Biolabs 公司; 混合纤维滤膜(水相, 直径 50 mm, 孔径分别为 1.2、0.8、0.45、0.22 和 0.1 μm), 海盐新东方塑化科技有限公司。砂芯过滤装置, 建湖县天一玻璃仪器厂; 无油隔膜真空泵, 长沙明杰仪器有限公司; 非接触式超声波破碎仪, Diagenode 公司; 高效液相色谱仪, Agilent 公司。宏基因组测序和 16S rRNA 基因 V3–V4 扩增子测序由北京诺禾致源科技股份有限公司完成。

## 1.3 样品处理

不同过滤方式: (1) 滤液过滤特别缓慢后才更换新的滤膜, 地下水样品 1、2 和 3 采用此种方式过滤; (2) 每过滤 1 L 水更换一张新膜, 地下水 4 采用此种方式过滤。

不同滤膜组合: 地下水样品 1、2 和 3 采用两种滤膜组合方式, 一种是直接用 0.1 μm 滤膜收集地下水中所有微生物, 分别编号为 GW10.1、GW20.1 和 GW30.1; 另一种组合使用 0.8 μm 和 0.1 μm 滤膜, 分别编号为 GW10.80.1、GW20.80.1 和 GW30.80.1。先用 0.8 μm 滤膜滤除细胞较大的微生物, 滤液过滤特别缓慢后更换新的滤膜, 每个样品过滤水样量为 100 L,

然后用 0.1 μm 滤膜收集水中剩余微生物。地下水样品 4 采用 8 种滤膜组合方式, 分别为 0.1 μm 滤膜单独, 1.2 μm 与 0.1 μm 滤膜组合, 0.8 μm 与 0.1 μm 滤膜组合, 0.45 μm 与 0.1 μm 滤膜组合, 0.22 μm 与 0.1 μm 滤膜组合, 0.8 μm、0.45 μm 与 0.1 μm 滤膜组合, 0.45 μm、0.45 μm 与 0.1 μm 滤膜组合, 以及 0.45、0.45、0.45 μm 与 0.1 μm 滤膜组合, 样品编号采用“GW4+滤膜孔径”的方式。过滤过程中, 除 0.1 μm 滤膜外, 其他滤膜不等滤膜孔塞满(滤液滤过速度未发生变化), 每 1 L 水更换一张新膜, 每种组合过滤水样为 25 L。过滤后的滤膜置于–80 °C 冰箱保存。

## 1.4 宏基因组测序

上步收集的滤膜分别无菌操作剪成小碎片, 采用十六烷基三甲基溴化铵(hexadecyl trimethyl ammonium bromide, CTAB)法提取样品基因组, 使用 Agilent 5400 检测浓度、完整性和纯度。基因组用非接触式超声波破碎仪(工作 15 s, 停 45 s, 共进行 6 个循环)随机打断至约 350 bp, 使用 NEBNext<sup>®</sup> Ultra DNA Library Preokit for Illumina 进行文库构建。库检合格后进行 Illumina PE150 测序(Illumina HiSeq 测序平台)。测序后, 使用 Readfq 对原始数据进行预处理, 去除低质量碱基超过 40 bp 的 reads, 去除 N 碱基超过 10 bp 的 reads, 去除与 adapter 超过 10 bp 重复的 reads 得到 clean data<sup>[20]</sup>。使用 MEGAHIT 软件对 clean reads 进行组装<sup>[21–22]</sup>, 并从 N 处打断获得不含 N 的 scaffigs<sup>[23–24]</sup>。使用 MetaGeneMark 对超过 500 bp 的 scaffigs 进行开放阅读框(open reading frame, ORF)预测<sup>[25]</sup>; 去除预测结果长度小于 100 nt 的信息后采用 CD-HIT 软件去冗余, 得到初始的非冗余基因集<sup>[26]</sup>。使用 Bowtie2 将 clean data 比对至初始 gene catalogue, 滤除掉比对数少于 3 个 reads

的基因获得用于后续分析的非冗余基因,并根据比对上的 reads 数和基因长度计算各基因的丰度<sup>[27]</sup>。DIAMOND 软件将 unigenes 与 Nr 数据库进行 BLASTp 同源比对<sup>[28]</sup>,并采用最低共同祖先(lowest common ancestor, LCA)算法获得各个样品在界门纲目科属种各个层级的基因数目表和丰度信息<sup>[29]</sup>。

### 1.5 16S rRNA 基因 V3-V4 扩增子测序

无菌水稀释样品 DNA 至 1 ng/ $\mu$ L 作为模板,选择 16S rRNA 基因 V3-V4 区引物 341F (5'-CCTAYGGGRBGCASCAG-3')和 806R (5'-G GACTACNNGGGTATCTAAT-3')<sup>[30]</sup>对细菌和古菌 V3-V4 区约 466 bp 序列进行扩增。PCR 反应体系: Phusion<sup>®</sup> High-Fidelity PCR Master Mix 15  $\mu$ L, 上、下游引物终浓度 0.2  $\mu$ mol/L, DNA 模板 10 ng。PCR 反应条件: 98  $^{\circ}$ C 1 min; 98  $^{\circ}$ C 10 s, 50  $^{\circ}$ C 30 s, 72  $^{\circ}$ C 30 s, 30 个循环; 72  $^{\circ}$ C 5 min。PCR 产物检测纯化后,使用 NEB Next<sup>®</sup> Ultra<sup>™</sup> II FS DNA PCR-free Library Prep Kit 进行文库构建,在 NovaSeq 6000 平台进行 PE 250 上机测序。测序后,去除 barcode 和引物序列,使用 FLASH 对 reads 进行拼接<sup>[31]</sup>,使用 fastp 软件获得高质量数据<sup>[32]</sup>。去除嵌合体<sup>[33]</sup>后得到的有效数据利用 QIME2 进行降噪(DADA2 方法)、物种注释(classify-sklearn 算法)、丰度计算和  $\alpha$ 、 $\beta$  多样性分析<sup>[34]</sup>。

宏基因组数据(登录号为 NMDC10018346)存储在国家微生物科学数据中心(National Microbiology Data Center, NMDC),链接为 <https://nmdc.cn/resource/genomics/metagenome/detail/NMDC10018346>。

## 2 结果与分析

### 2.1 地下水样品中 CPR 和 DPANN 多样性

使用 0.1  $\mu$ m 滤膜收集地下水中微生物后进

行宏基因组测序。4 个样品(GW10.1-GW40.1) 组装平均获得 188 303 条超过 500 bp 的 scaftigs,其中样品 GW40.1 最多,样品 GW30.1 最少(表 1)。Scaftigs 平均长度在 1 075-1 762 bp 之间,总长平均为 274 365 379 bp。4 个样品 ORF 预测分别得到 423 967、244 416、419 888 和 423 967 个基因。与 Nr 数据库中抽提出的细菌、真菌和古菌序列进行比对并获得序列的物种注释信息后发现,不论从物种多样性还是丰度上均说明地下水中存在丰富的细菌物种资源(表 2)。

4 个地下水样品中变形菌门(*Proteobacteria*, 现更名为 *Pseudomonadota*<sup>[35]</sup>)、厚壁菌门(*Firmicutes*, 现更名为 *Bacillota*<sup>[35]</sup>)和放线菌门(*Actinobacteria*, 现更名为 *Actinomycetota*<sup>[35]</sup>)相对丰度均超过 1%,其中以变形菌门相对丰度最高,分别为 73.60%、89.17%、64.69%和 46.72%。其他相对丰度超过 1%的菌门还包括 GW30.1 中的异常球菌-栖热菌门(*Deinococcus-Thermus*, 现更名为 *Deinococcota*<sup>[35]</sup>)和 GW40.1 中的拟杆菌门(*Bacteroidetes*, 现更名为 *Bacteroidota*<sup>[35]</sup>)。从门水平上分析,地下水含有种类丰富的 CPR 和 DPANN 菌门,CPR 类群检出 33-64 个菌门,占检出细菌门种类的 40.2%-47.4%; DPANN 检出 1-7 个菌门,占检出古菌门种类的 25.0%-45.5% (图 1A)。不同地下水在检出的 CPR 菌门种类和丰度方面存在较大的差异,GW10.1 中相对丰度前 3 的 CPR 菌门分别为 *Candidatus Adlerbacteria*、*Ca. Kaiserbacteria* 和 *Ca. Staskawiczbacteria*, GW20.1 中相对丰度前 3 的 CPR 菌门分别为 *Ca. Gracilibacteria*、*Ca. Peregrinibacteria* 和 *Ca. Magasanikbacteria*, GW30.1 中相对丰度前 3 的 CPR 菌门分别为 *Ca. Parcubacteria*、*Saccharibacteria* 和 *Ca. Kaiserbacteria*, GW40.1 中相对丰度前 3 的 CPR 菌门分别为 *Saccharibacteria*、*Ca. Levybacteria* 和 *Ca. Wildermuthbacteria* (图 1B)。

表 1 宏基因组测序及 16S rRNA 基因 V3–V4 区扩增子测序数据

Table 1 Metagenomic sequencing and 16S rRNA gene V3–V4 amplicon sequencing data

Sequencing	Sample	Sample ID	BaseNum (bp)	SeqNum	MeanLen (bp)	MaxLen (bp)	NMDC accession No.
Metagenomic sequencing	Groundwater 1	GW10.1	299 326 439	202 993	1 475	277 085	NMDC40030136
	Groundwater 1	GW10.80.1	176 599 666	109 443	1 614	474 139	NMDC40030137
	Groundwater 2	GW20.1	305 654 347	189 917	1 609	574 941	NMDC40030138
	Groundwater 2	GW20.80.1	339 092 579	214 334	1 582	1 325 421	NMDC40030139
	Groundwater 3	GW30.1	269 698 075	153 100	1 762	494 122	NMDC40030140
	Groundwater 3	GW30.80.1	360 120 764	235 997	1 526	1 149 745	NMDC40030141
	Groundwater 4	GW40.1	222 782 656	207 203	1 075	250 836	NMDC40030142
	Groundwater 4	GW41.20.1	332 463 971	277 628	1 198	490 005	NMDC40030143
	Groundwater 4	GW40.80.1	247 301 282	278 064	889	104 695	NMDC40030144
	Groundwater 4	GW430.450.1	206 851 879	183 716	1 126	495 059	NMDC40030145
16S rRNA gene V3–V4 amplicon sequencing	Groundwater 1	GW10.116S	28 312 363	67 869	417		NMDC40030146
	Groundwater 2	GW20.116S	34 935 159	84 661	413		NMDC40030147
	Groundwater 3	GW30.116S	30 597 286	73 049	419		NMDC40030148
	Groundwater 4	GW40.116S	50 986 351	124 059	410		NMDC40030149

Sample ID: 样品名称; BaseNum (bp): 组装得 scaffigs/tags 序列的总长; SeqNum: 组装得到的 scaffigs/tags 序列总条数; MeanLen (bp): Scaffigs/Tags 序列的平均长度; MaxLen (bp): 组装得到的最长 scaffigs/tags 序列的长度值

Sample ID: The sample name; BaseNum (bp): The total length of the assembled scaffigs/tags sequence; SeqNum: The total number of the assembled scaffigs/tags sequences; MeanLen (bp): The average length of the scaffigs/tags sequence; MaxLen (bp): The longest length of the assembled scaffigs/tags sequence.

表 2 地下水样品微生物门水平分布情况

Table 2 Microorganism distribution in phylum level in groundwater samples

Kingdom	GW10.1		GW20.1		GW30.1		GW40.1	
	Phylum number	Relative abundance (%)	Phylum number	Relative abundance (%)	Phylum number	Relative abundance (%)	Phylum number	Relative abundance (%)
Bacteria	134	89.02	82	93.22	119	86.74	136	83.91
Archaea	14	1.64	4	0.01	11	0.04	16	1.07
Fungi	5	0.01	4	0.01	8	0.04	8	0.12

不同地下水在检出的 DPANN 菌门种类上有所差异, 但从相对丰度方面分析, 其中的 *Ca. Woesearchaeota* 在 4 个地下水 DPANN 中的占比最高, 分别占检出 DPANN 量的 69.9%、100.00%、76.81% 和 55.44% (图 1C)。

地下水样品中虽然存在多样性极高的 CPR 和 DPANN 物种资源, 但是其相对丰度较低, 分别在 0.17%–1.67% 和 0.000 93%–0.071 00% 之

间(图 1A)。其中, 地下水样品 GW40.1 在检出 CPR 和 DPANN 种类和丰度方面均具有较大优势, CPR 共检出 64 个菌门, 总的相对丰度为 1.67%; DPANN 共检出 7 个菌门, 总的相对丰度为 0.071% (图 1)。地下水样品 GW40.1 检出的 CPR 菌门中, 相对丰度超过 0.1% 的有 2 个菌门, 分别是 *Saccharibacteria* 和 *Ca. Levybacteria*, 相对丰度在 0.01%–0.1% 的有 28 个菌门, 分别

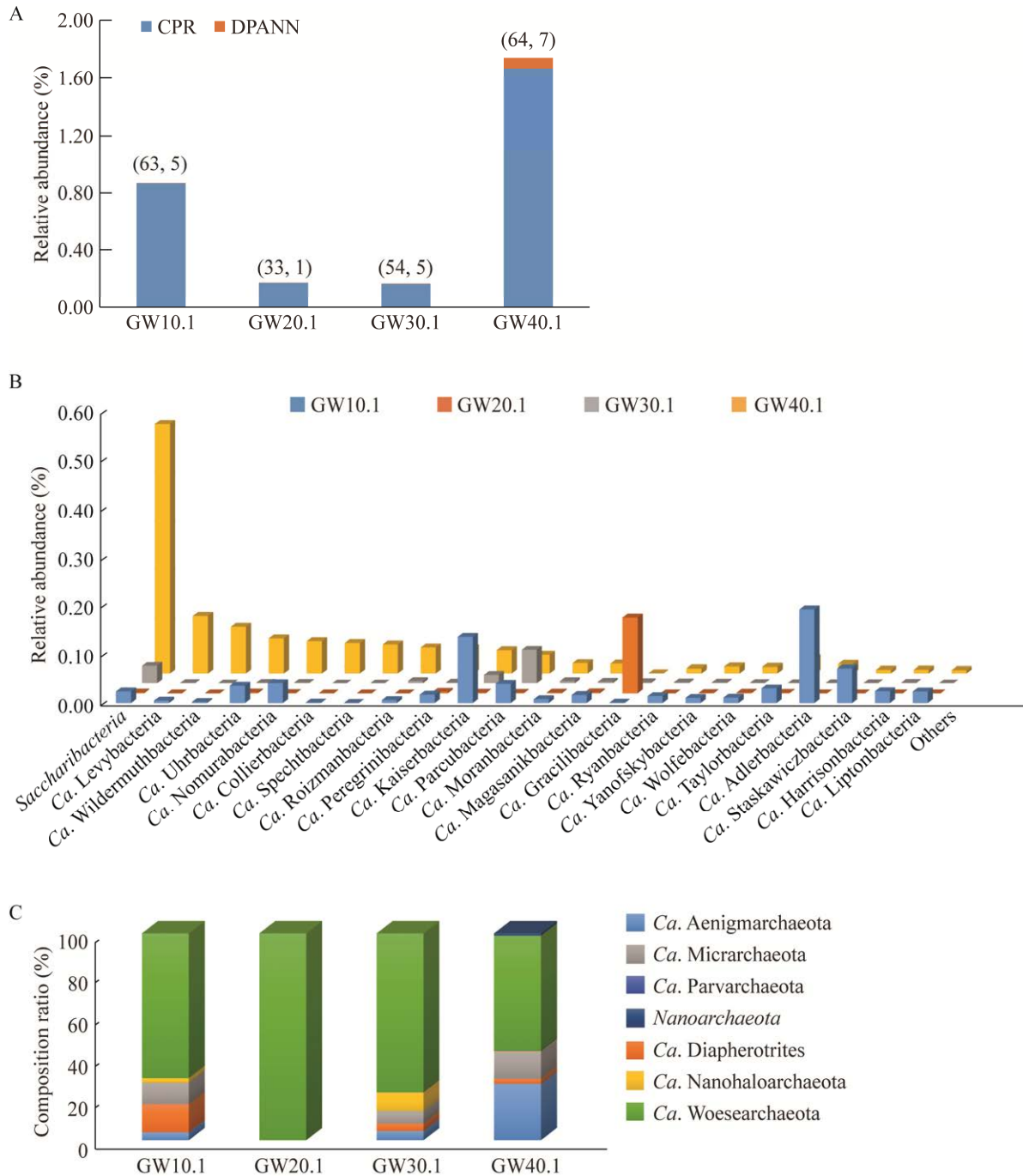


图 1 呼和浩特市周边地下水中 CPR 和 DPANN 含量 A: 不同地下水样品中 CPR 和 DPANN 相对丰度和检出的门水平数目. B: 不同地下水样品中丰度前 10 的 CPR 菌门相对丰度柱形图. C: 不同地下水样品中 DPANN 的结构组成

Figure 1 CPR and DPANN contents in groundwater around Hohhot City. A: CPR and DPANN relative abundance and the detected phylum numbers in different groundwater samples. B: Relative abundance bar charts of the top 10 CPR phyla in different groundwater samples. C: Composition ratio of DPANN in different groundwater samples.

是 *Ca. Wildermuthbacteria*、*Ca. Uhrbacteria*、*Ca. Nomurabacteria*、*Ca. Collierbacteria*、*Ca. Spechtbacteria*、*Ca. Roizmanbacteria*、*Ca. Peregrinibacteria*、*Ca. Kaiserbacteria*、*Ca. Sungbacteria*、*Parcubacteria*、*Ca. Taylorbacteria*、*Ca. Neelsonbacteria*、*Ca. Gottesmanbacteria*、*Ca. Daviesbacteria*、*Ca. Giovannonibacteria*、*Ca. Moranbacteria*、*Ca. Magasanikbacteria*、*Ca. Adlerbacteria*、*Ca. Woesebacteria*、*Ca. Azambacteria*、*Ca. Doudnabacteria*、*Ca. Pacebacteria*、*Ca. Yanofskybacteria*、*Ca. Wolfbacteria*、*Ca. Yonathbacteria*、*Ca. Niyogibacteria*、*Ca. Falkowbacteria* 和 *Ca. Ryanbacteria*。然而, 即使是其中丰度最高的 *Saccharibacteria* 菌门, 相对丰度也仅有 0.51%, 仅排在 11 位。地下水样品 GW40.1 检出的 DPANN 菌门中, 相对丰度超过 0.01% 的仅有 2 个菌门, 分别是 *Ca. Woesearchaeota* 和 *Ca. Aenigmarchaeota*。

## 2.2 不同测序方法对 CPR 和 DPANN 检出的影响

如表 1 所示, 4 个地下水样品 16S rRNA 基因 V3-V4 区扩增子测序共产生 349 638 条有效序列, 平均 87 409 条, 有效序列均超过 4 万条, 其中 GW40.116S 最多, GW10.116S 最少。采用 DADA2 方法去重后得到扩增序列变体(amplicon sequence variants, ASVs)分别为 870、1 285、651 和 1 444 个, 其中特异 ASVs 分别为 593、1 168、385 和 1 282 个(图 2)。 $\alpha$  多样性分析中, Chao1 用来估计样品中含有的 OTU 数目; Coverage 反映样品文库覆盖率, 数值越接近 1 表示样品中序列未被测出的概率越低; Shannon 和 Simpson 指数用来评估样品中微生物多样性, Shannon 越大、Simpson 越低说明样品中微生物群落多样性越高(表 3)。16S

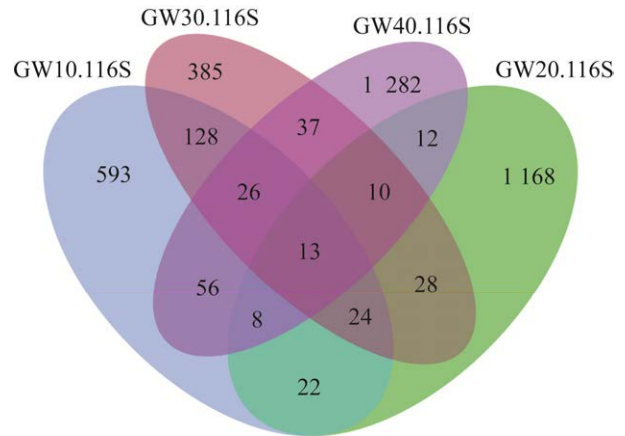


图 2 16S rRNA 基因 V3-V4 扩增子测序韦恩图  
Figure 2 Venn diagram of 16S rRNA gene V3-V4 amplicon sequencing.

表 3 16S rRNA 基因 V3-V4 区扩增子测序  $\alpha$  多样性分析

Table 3 Alpha diversity analysis of the 16S rRNA gene V3-V4 amplicon sequencing samples

Sample	Shannon	Chao1	Good coverage	Simpson
GW10.116S	6.444	878.041	0.999	0.950
GW20.116S	6.642	1 323.521	0.997	0.960
GW30.116S	6.617	662.000	0.999	0.972
GW40.116S	5.914	1 491.962	0.997	0.809

rRNA 基因 V3-V4 区扩增子测序结果反映出 4 个地下水样品具有相近的微生物菌群结构。进一步比较宏基因组测序和 16S rRNA 基因 V3-V4 扩增子测序这 2 种方法对 CPR 和 DPANN 检出的影响。结果如图 3 所示, 4 个地下水样品富含 CPR 和 DPANN 菌株, 但 16S rRNA 基因 V3-V4 扩增子测序在这 4 个地下水样品中均未检出 CPR 菌门, DPANN 中仅检出 *Nanoarchaeota*。

## 2.3 滤膜处理对 CPR 和 DPANN 菌株富集的影响

多样性研究显示, 4 个地下水样品均存在种类丰富的 CPR 和 DPANN 物种资源, 然而较低的存在丰度和营共生生活、难培养的特性制



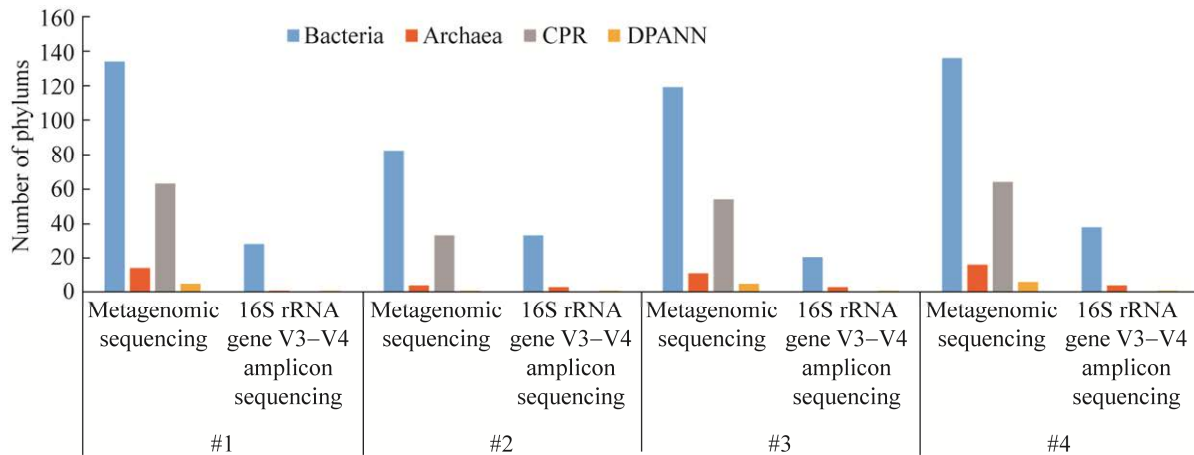


图3 宏基因组测序与16S rRNA基因V3-V4扩增子测序对CPR和DPANN检出的影响

Figure 3 Effects of metagenomic sequencing and 16S rRNA gene V3-V4 amplicon sequencing on detection of CPR and DPANN.

约着对这两类菌的深入研究。考虑到 CPR 和 DPANN 菌株具有细胞体积较普通细菌和古菌更小的共同特征, 本文研究了不同滤膜孔径和不同过滤方式对 CPR 和 DPANN 的富集作用。地下水样品 1、2 和 3 采用不及时更换滤膜的方式比较了单独使用 0.1  $\mu\text{m}$  滤膜和组合使用 0.8  $\mu\text{m}$  和 0.1  $\mu\text{m}$  滤膜的过滤富集效果。如图 4A、4B 所示, 在此种过滤方式下使用 0.8  $\mu\text{m}$  进行预过滤并未达到去除部分细胞体积较大菌体、富集 CPR 和 DPANN 的目的。组合使用 0.8  $\mu\text{m}$  和 0.1  $\mu\text{m}$  滤膜过滤后, 得到的 CPR 和 DPANN 相对丰度均比单独使用 0.1  $\mu\text{m}$  滤膜过滤要低。地下水样品 4 采用及时更换滤膜的方式(每隔 1 L 水更换一张新滤膜), 比较了不同滤膜组合对 CPR 和 DPANN 富集的影响。0.45  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合, 0.22  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合, 0.8  $\mu\text{m}$ 、0.45  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合, 以及 0.45  $\mu\text{m}$ 、0.45  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合由于基因组提取及文库构建原因未成功。比较成功的 4 组发现, 1.2  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合, 0.8  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合以及 0.45、0.45、0.45  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤

膜组合得到的 CPR 和 DPANN 相对丰度均比单独使用 0.1  $\mu\text{m}$  滤膜过滤要高(图 4C), 说明此种过滤方式下, 大孔径滤膜能去除部分细胞体积较大菌体, 实现富集 CPR 和 DPANN 的目的。其中 1.2  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合效果最好, CPR 相对丰度由 1.67% 提高到 13.33%, 提高 8.0 倍; DPANN 相对丰度由 0.071% 提高到 0.58%, 提高 8.2 倍。同时 1.2  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合对所有 CPR 和 DPANN 菌株均有较好的富集效果, 地下水样品 4 共检出 64 个 CPR 菌门, 处理后相对丰度提高了 2.1–16.8 倍; 共检出 7 个 DPANN 菌门, 处理后相对丰度提高 5.0–10.0 倍(表 4)。

### 3 讨论

CPR 细菌和 DPANN 古菌是与绝大多数已知细菌和古菌在系统发育地位、基因组及细胞大小和生理代谢等方面具有显著差异的独特类群。进化树显示两类群分别位于细菌域和古菌域外侧, 相互毗邻<sup>[9]</sup>。两类群具有许多共性特征, 如细胞微小、基因组精简、缺少部分核心

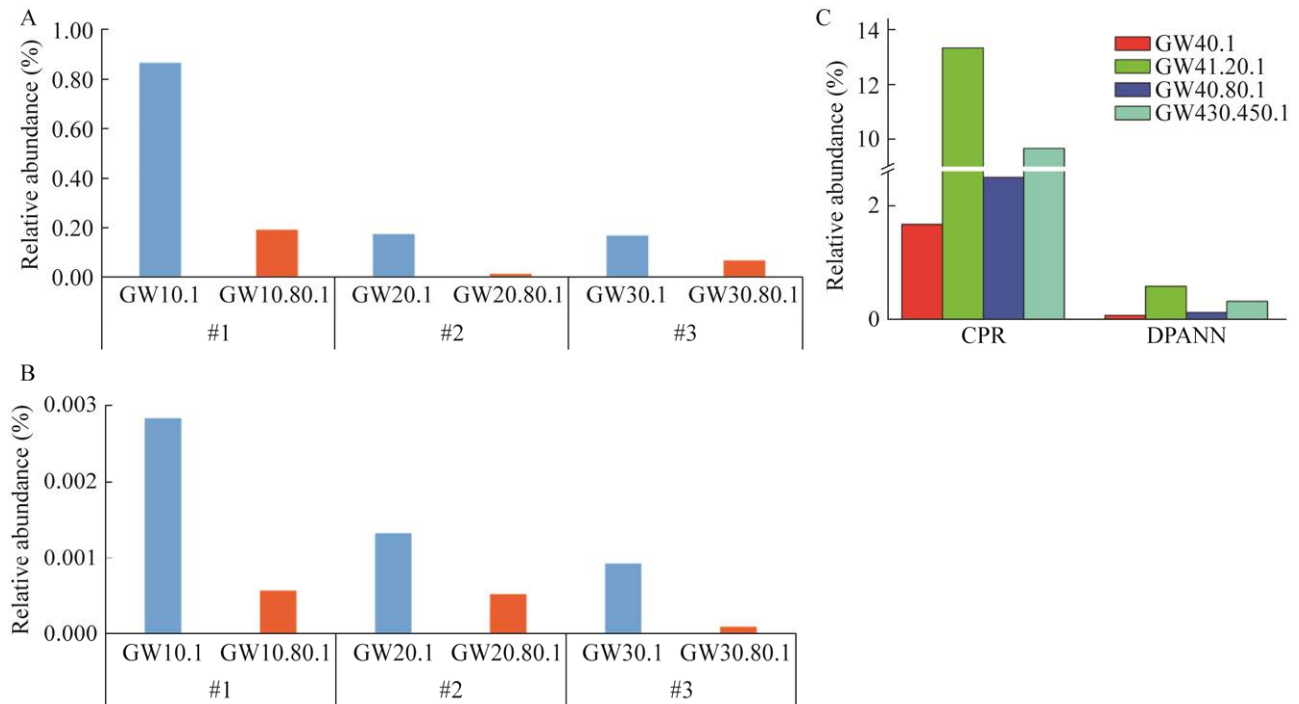


图 4 不同过滤方式对 CPR 和 DPANN 菌株的富集作用 A: 地下水样品 1、2 和 3 采用不同过滤处理对 CPR 菌株收集的影响. B: 地下水样品 1、2 和 3 采用不同过滤处理对 DPANN 菌株收集的影响. C: 地下水样品 4 采用不同过滤处理对 CPR 和 DPANN 菌株收集的影响

Figure 4 Enrichment of CPR and DPANN strains by different filtration methods. A: The effects of different filtration treatments on the collection of CPR strains in groundwater samples 1, 2 and 3. B: The effects of different filtration treatments on the collection of DPANN strains in groundwater samples 1, 2 and 3. C: The effects of different filtration treatments on the collection of CPR and DPANN strains in groundwater samples 4.

代谢能力、厌氧/耐氧、难培养和营共生生活等。已知 CPR 和 DPANN 类群庞大且存在广泛, 在地下水、湖泊、极端环境、海洋、动物和人体口腔等各种生境都有发现。然而, 由于发现较晚, 目前对于 CPR 和 DPANN 的研究仍处于起步阶段, 在不同生境中多样性的研究还较少, 生态功能尚属未知。已有研究显示, 很多 CPR 和 DPANN 细胞能够通过  $0.2 \mu\text{m}$  滤膜<sup>[36-38]</sup>。Brown 等<sup>[3]</sup>利用  $1.2$ 、 $0.2$  和  $0.1 \mu\text{m}$  滤膜组合过滤地下水, 发现  $0.1 \mu\text{m}$  滤膜富集的 CPR 细菌相对丰度要比  $0.2 \mu\text{m}$  滤膜多 50% 以上。本研究利用  $0.1 \mu\text{m}$  滤膜收集呼和浩特市周边 4 个不同地

下水样品中微生物, 宏基因组测序分析其中 CPR 和 DPANN 多样性, 发现地下水中存在丰富的 CPR 和 DPANN 物种资源, 最多的地下水样品 4 中检出了 64 个 CPR 菌门和 7 个 DPANN 菌门, 表明目前已知的 CPR 菌门中有超过 80% 被检出, 已知的 DPANN 菌门中有超过 50% 被检出。同时发现不同的地下水在 CPR 和 DPANN 存在种类和丰度方面均有较大差异, 这与 He 等<sup>[4]</sup>和 Probst 等<sup>[39]</sup>的结论一致。

研究表明, 很多 CPR 类群菌株 16S rRNA 和 23S rRNA 基因中存在内含子, 且在保守区和可变区都被发现<sup>[3]</sup>; DPANN 超门中 *Ca.*

表 4 不同滤膜组合对地下水样品 4 中 CPR 和 DPANN 富集的影响

Table 4 Effects of different membrane combinations on CPR and DPANN enrichment in groundwater sample 4

Group	Phylum/Code name	Relative abundance (%)			
		GW40.1	GW41.20.1	GW40.80.1	GW430.450.1
CPR	<i>Saccharibacteria</i>	0.514 18	2.917 30	0.830 05	1.073 81
	<i>Ca. Collierbacteria</i>	0.062 25	1.047 36	0.072 07	0.683 00
	<i>Ca. Levybacteria</i>	0.117 83	0.967 66	0.118 16	1.811 31
	<i>Ca. Uhrbacteria</i>	0.071 61	0.867 46	0.108 16	0.234 78
	<i>Ca. Wildermuthbacteria</i>	0.095 54	0.859 46	0.150 23	0.843 72
	<i>Ca. Nomurabacteria</i>	0.066 04	0.639 80	0.097 27	0.489 64
	<i>Ca. Spechtbacteria</i>	0.059 15	0.586 97	0.090 38	0.283 27
	<i>Ca. Sungbacteria</i>	0.047 01	0.448 87	0.090 87	0.378 94
	<i>Ca. Peregrinibacteria</i>	0.051 25	0.396 14	0.055 61	0.142 65
	<i>Ca. Roizmanbacteria</i>	0.052 99	0.355 14	0.062 37	0.425 47
	<i>Ca. Nealsobacteria</i>	0.028 44	0.213 74	0.040 97	0.159 40
	<i>Ca. Adlerbacteria</i>	0.019 43	0.190 13	0.043 54	0.333 41
	<i>Ca. Gottesmanbacteria</i>	0.026 99	0.180 90	0.023 18	0.121 66
	<i>Ca. Magasanikbacteria</i>	0.020 02	0.177 06	0.027 49	0.084 66
	<i>Ca. Taylorbacteria</i>	0.030 48	0.176 44	0.062 57	0.063 42
	<i>Ca. Daviesbacteria</i>	0.021 34	0.167 45	0.026 06	0.272 38
	<i>Ca. Moranbacteria</i>	0.021 05	0.160 95	0.024 36	0.070 77
	<i>Ca. Azambacteria</i>	0.015 89	0.141 28	0.024 73	0.072 03
	<i>Ca. Woesebacteria</i>	0.016 82	0.141 08	0.024 68	0.124 70
	<i>Ca. Doudnabacteria</i>	0.014 95	0.111 88	0.018 91	0.043 82
	<i>Ca. Yanofskybacteria</i>	0.014 27	0.104 04	0.019 87	0.072 51
	<i>Ca. Pacebacteria</i>	0.014 84	0.097 54	0.019 60	0.078 28
	<i>Ca. Niyogibacteria</i>	0.011 97	0.090 89	0.020 02	0.055 96
	<i>Ca. Wolfbacteria</i>	0.013 47	0.090 28	0.019 82	0.104 69
	<i>Ca. Yonathbacteria</i>	0.012 94	0.087 20	0.022 48	0.026 90
	<i>Ca. Falkowbacteria</i>	0.010 56	0.085 29	0.013 59	0.025 81
	<i>Ca. Ryanbacteria</i>	0.010 29	0.078 23	0.015 54	0.042 49
	<i>Ca. Buchananbacteria</i>	0.006 92	0.069 80	0.010 36	0.033 97
	<i>Ca. Portnoybacteria</i>	0.008 31	0.065 48	0.011 92	0.037 29
	<i>Ca. Zambryskibacteria</i>	0.008 01	0.064 05	0.011 61	0.074 20
	<i>Ca. Harrisonbacteria</i>	0.007 61	0.058 32	0.011 49	0.051 22
	<i>Ca. Beckwithbacteria</i>	0.004 43	0.057 16	0.004 97	0.040 86
	<i>Ca. Staskawiczbacteria</i>	0.007 33	0.055 48	0.009 45	0.051 32
	<i>Ca. Amesbacteria</i>	0.005 35	0.048 42	0.004 93	0.036 03
	<i>Ca. Curtissbacteria</i>	0.006 76	0.045 19	0.007 06	0.088 68
	<i>Ca. Kerfeldbacteria</i>	0.005 08	0.044 85	0.006 20	0.012 56
	<i>Ca. Liptonbacteria</i>	0.006 53	0.044 54	0.009 69	0.021 15
	<i>Ca. Shapirobacteria</i>	0.003 72	0.040 12	0.004 91	0.034 03
	<i>Ca. Lloydbacteria</i>	0.005 34	0.038 61	0.008 05	0.030 70

(待续)

(续表 4)

Group	Phylum/Code name	Relative abundance (%)			
		GW40.1	GW41.20.1	GW40.80.1	GW430.450.1
	<i>Ca. Berkelbacteria</i>	0.004 18	0.028 96	0.005 60	0.013 66
	<i>Ca. Komeilibacteria</i>	0.003 64	0.026 88	0.004 24	0.007 19
	<i>Ca. Terrybacteria</i>	0.003 03	0.024 15	0.005 48	0.013 09
	<i>Ca. Vogelbacteria</i>	0.003 54	0.024 01	0.005 66	0.045 55
	<i>Ca. Chisholmbacteria</i>	0.001 87	0.021 23	0.001 79	0.012 92
	<i>Ca. Campbellbacteria</i>	0.002 81	0.018 96	0.004 48	0.012 87
	<i>Ca. Woykebacteria</i>	0.002 74	0.018 66	0.002 80	0.020 74
	<i>Ca. Kuenenbacteria</i>	0.003 58	0.015 82	0.004 60	0.009 43
	<i>Ca. Colwellbacteria</i>	0.001 98	0.014 57	0.002 46	0.008 53
	<i>Ca. Tagabacteria</i>	0.002 11	0.014 16	0.002 81	0.009 26
	CPR2	0.001 42	0.011 74	0.001 70	0.006 79
	<i>Ca. Abawacabacteria</i>	0.001 13	0.010 85	0.001 44	0.004 56
	<i>Ca. Blackburnbacteria</i>	0.002 35	0.009 68	0.002 42	0.014 80
	<i>Ca. Andersenbacteria</i>	0.000 74	0.007 05	0.002 11	0.002 88
	CPR3	0.000 91	0.006 65	0.001 33	0.003 45
	<i>Ca. Jacksonbacteria</i>	0.000 51	0.003 33	0.000 69	0.001 58
	CPR1	0.001 21	0.002 56	0.001 00	0.001 03
	<i>Ca. Gracilibacteria</i>	0.000 53	0.002 47	0.000 51	0.001 14
	<i>Ca. Brennerbacteria</i>	0.000 18	0.001 96	0.000 55	0.001 46
	<i>Ca. Veblenbacteria</i>	0.000 16	0.001 28	0.000 25	0.000 32
	<i>Ca. Giovannonibacteria</i>	0.021 11	0.180 36	0.035 78	0.170 17
	<i>Ca. Jorgensenbacteria</i>	0.004 17	0.032 72	0.006 40	0.018 65
	<i>Ca. Kaiserbacteria</i>	0.047 32	0.456 83	0.095 80	0.381 52
	<i>Microgenomates</i> (OP11)	0.002 20	0.019 25	0.003 66	0.014 64
	<i>Parcubacteria</i> (OD1)	0.037 98	0.364 14	0.086 87	0.255 69
DPANN	<i>Ca. Aenigmarchaeota</i>	0.019 58	0.127 27	0.033 68	0.113 06
	<i>Ca. Diapherotrites</i>	0.001 65	0.015 60	0.002 51	0.008 74
	<i>Ca. Micrarchaeota</i>	0.009 37	0.065 39	0.015 65	0.041 21
	<i>Ca. Nanohaloarchaeota</i>	0.000 21	0.001 06	0.000 69	0.000 51
	<i>Ca. Parvarchaeota</i>	0.000 20	0.001 33	0.000 33	0.001 08
	<i>Ca. Woesearchaeota</i>	0.039 60	0.372 52	0.064 13	0.151 51
	<i>Nanoarchaeota</i>	0.000 12	0.001 20	0.000 35	0.000 78

*Aenigmarchaeota*、*Ca. Micrarchaeota*、*Ca. Pacearchaeota* 和 *Ca. Woesearchaeota* 16S rRNA 基因中发现有蛋白编码基因的存在<sup>[1]</sup>。同时, CPR 和 DPANN 类群在 16S rRNA 基因序列层面高度可变, 很多是在保守区, 容易造成 16S rRNA 基因通用引物不匹配的现象<sup>[1]</sup>。因此本研究比较了宏基因组测序和 16S rRNA 基因

V3-V4 区扩增子测序两种方法对于 CPR 和 DPANN 检出结果的影响。其中 16S rRNA 基因 V3-V4 区扩增子测序分析采用 DADA2 方法进行降噪去重, 相当于 100% 相似度聚类。DADA2 方法比传统的 OTU 方法更敏感和特异, 输出的假序列更少且能够检测到 OTU 方法遗漏的真实生物变异<sup>[40]</sup>。DADA2 方法去重

后得到的 ASVs 也称特征序列, 对应于 OTU 代表序列, ASVs 替代 OTU 能够提高标记基因数据分析的全面性、准确性和可重复性<sup>[41]</sup>。本研究发现 16S rRNA 基因 V3-V4 区扩增子测序很难检测出样品中的 CPR 和 DPANN。4 个地下水仅检出 DPANN 超门中的 *Nanoarchaeota* 门。考虑到目前对环境样品微生物多样性的研究多采用 16S rRNA 基因 V3-V4 区扩增子测序的方法, 遗漏了 CPR 和 DPANN 的信息。因此, 有必要对 CPR 和 DPANN 存在生境和多样性开展系统研究。

厌氧/耐氧、难培养和营共生生活的特点导致 CPR 和 DPANN 缺少分离培养菌株, 基因功能难验证, 天然产物资源难以挖掘。目前 CPR 只有 *Saccharibacteria*<sup>[8,36,42-44]</sup>、*Ca. Absconditabacteria*<sup>[42]</sup>、*Ca. Gracilibacteria*<sup>[45]</sup> 和 *Ca. Yanofskybacteria*<sup>[46-47]</sup> 的部分菌株实现了实验室培养; DPANN 只有 *Nanoarchaeota*<sup>[37,48]</sup>、*Ca. Nanohaloarchaeota*<sup>[38,49]</sup> 和 *Ca. Micrarchaeota*<sup>[50]</sup> 的部分菌株实现了实验室培养, 且均与其他宿主菌共生生活。对于 CPR 和 DPANN 物种资源的挖掘更多的是通过宏基因组深度测序后进行 binning 组装, 选择组装后完整性较高和污染度较低的宏基因组组装基因组 (metagenome-assembled genomes, MAGs) 进行物种注释和生物信息学分析。然而, CPR 和 DPANN 在生境中的存在丰度普遍较低, 即使是多样性和丰度具有明显优势的地下水样品中 CPR 和 DPANN 仍较常规细菌和古菌丰度低很多, 影响宏基因组深度测序获得 CPR 和 DPANN 的 MAGs 数量和质量。本研究检测的 4 个地下水样品中, CPR 丰度最低的 GW30.1 仅为 0.17%, 最高的 GW40.1 也只有 1.67%, DPANN 丰度最低的 GW30.1 仅为 0.000 93%, 最高的 GW40.1 也只有 0.071%。提高 CPR 和 DPANN 在样品中的相对丰度, 同时

在减少其他细菌和古菌干扰的前提下尽可能多地收集 CPR 和 DPANN 菌株, 是后续通过深度测序获得高质量 CPR 和 DPANN 菌株基因组、通过文库构建筛选 CPR 和 DPANN 中天然产物资源和基因资源、通过可培养研究分离 CPR 和 DPANN 菌株的重要基础。本研究比较了不及时更换滤膜和及时更换滤膜两种过滤方式及不同孔径滤膜组合对于 CPR 和 DPANN 的富集效果。结果显示, 3 个地下水样品不及时更换 0.8  $\mu\text{m}$  滤膜 (GW10.80.1、GW20.80.1 和 GW30.80.1) 均比仅用 0.1  $\mu\text{m}$  滤膜 (GW10.1、GW20.1 和 GW30.1) 收集的 CPR 和 DPANN 相对丰度低, 说明应在过滤过程中经常更换以滤除细胞较大微生物为目的的滤膜, 防止滤膜孔堵塞, 将游离的 CPR 和 DPANN 细胞截流, 从而达不到富集的目的。地下水样品 4 及时更换滤膜, 也就是不等滤膜孔塞满, 滤液过滤速度未发生变化时, 每 1 L 水就更换一张新膜。利用此种过滤方式, 发现 3 种滤膜组合 (GW41.20.1、GW40.80.1 和 GW430.450.1) 得到的 CPR 和 DPANN 相对丰度均比单独使用 0.1  $\mu\text{m}$  滤膜 (GW410.1) 过滤要高。其中 1.2  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合效果最好, CPR 和 DPANN 相对丰度分别提高 8.0 倍和 8.2 倍。一方面, 1.2  $\mu\text{m}$  孔径滤膜能截留并去除细胞体积较大的菌体; 另一方面, CPR 和 DPANN 菌株多营共生生活, 1.2  $\mu\text{m}$  孔径滤膜能允许部分共生菌及共生的 CPR 和 DPANN 细胞通过。Castelle 等<sup>[17]</sup> 的研究也体现了这一点, 他们利用 1.2  $\mu\text{m}$  滤膜预过滤, 0.2  $\mu\text{m}$  和 0.1  $\mu\text{m}$  滤膜收集菌体, 发现能有效富集 DPANN 菌株。

## 4 结论

本研究采用 0.1  $\mu\text{m}$  滤膜收集菌体和宏基因组测序的方法分析了呼和浩特市周边 4 个不同地下水中 CPR 和 DPANN 的多样性, 发现地下

水中存在种类丰富但相对丰度较低的 CPR 和 DPANN 物种资源。同时,不同地下水中的 CPR 和 DPANN 在种类和丰度方面均有一定的差异。研究工作比较了宏基因组测序与 16S rRNA 基因扩增子测序对 CPR 和 DPANN 检出的影响,发现常规的 16S rRNA 基因 V3-V4 区扩增子测序很难检出 CPR 和 DPANN 物种信息,后续工作有必要对 CPR 和 DPANN 存在生境和多样性开展系统研究。本次实验还研究了不同过滤方式和不同滤膜组合对 CPR 和 DPANN 富集的作用,发现在及时更换滤膜的前提下,采用不同孔径滤膜组合过滤地下水可以有效富集 CPR 和 DPANN,其中以 1.2  $\mu\text{m}$  与 0.1  $\mu\text{m}$  滤膜组合效果最好,过滤后 CPR 和 DPANN 相对丰度分别提高到 13.33% 和 0.58%。本研究为下一步挖掘 CPR 和 DPANN 物种资源、基因资源和天然产物资源,以及对 CPR 和 DPANN 菌株可培养研究提供了参考。同时,后续工作可以尝试在滤膜过滤的基础上引入其他方法进一步富集 CPR 和 DPANN 细胞,减少其他细菌和古菌的干扰。

## REFERENCES

- [1] CASTELLE CJ, BROWN CT, ANANTHARAMAN K, PROBST AJ, HUANG RH, BANFIELD JF. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations[J]. *Nature Reviews Microbiology*, 2018, 16(10): 629-645.
- [2] RINKE C, SCHWIENSTEK P, SCZYRBA A, IVANOVA NN, ANDERSON IJ, CHENG JF, DARLING A, MALFATTI S, SWAN BK, GIES EA, DODSWORTH JA, HEDLUND BP, TSIAMIS G, SIEVERT SM, LIU WT, EISEN JA, HALLAM SJ, KYRPIDES NC, STEPANAUSKAS R, RUBIN EM, et al. Insights into the phylogeny and coding potential of microbial dark matter[J]. *Nature*, 2013, 499(7459): 431-437.
- [3] BROWN CT, HUG LA, THOMAS BC, SHARON I, CASTELLE CJ, SINGH A, WILKINS MJ, WRIGHTON KC, WILLIAMS KH, BANFIELD JF. Unusual biology across a group comprising more than 15% of domain Bacteria[J]. *Nature*, 2015, 523(7559): 208-211.
- [4] HE C, KEREN R, WHITTAKER ML, FARAG IF, DOUDNA JA, CATE JHD, BANFIELD JF. Genome-resolved metagenomics reveals site-specific diversity of episymbiotic CPR bacteria and DPANN archaea in groundwater ecosystems[J]. *Nature Microbiology*, 2021, 6(3): 354-365.
- [5] NICOLAS AM, JAFFE AL, NUCCIO EE, TAGA ME, FIRESTONE MK, BANFIELD JF. Soil candidate phyla radiation bacteria encode components of aerobic metabolism and co-occur with nanoarchaea in the rare biosphere of rhizosphere grassland communities[J]. *mSystems*, 2021, 6(4): e0120520.
- [6] CAI RN, ZHANG J, LIU R, SUN CM. Metagenomic insights into the metabolic and ecological functions of abundant deep-sea hydrothermal vent DPANN archaea[J]. *Applied and Environmental Microbiology*, 2021, 87(9): e03009-e03020.
- [7] DOMBROWSKI N, LEE JH, WILLIAMS TA, OFFRE P, SPANG A. Genomic diversity, lifestyles and evolutionary origins of DPANN archaea[J]. *FEMS Microbiology Letters*, 2019, 366(2): fnz008.
- [8] HE XS, MCLEAN JS, EDLUND A, YOOSEPH S, HALL AP, LIU SY, DORRESTEIN PC, ESQUENAZI E, HUNTER RC, CHENG GH, NELSON KE, LUX R, SHI WY. Cultivation of a human-associated TM7 phylotype reveals a reduced genome and epibiotic parasitic lifestyle[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2015, 112(1): 244-249.
- [9] 陶晔, 邢鹏. 候选门级辐射类群(CPR)细菌研究进展[J]. *微生物学报*, 2020, 60(6): 1284-1303.
- TAO Y, XING P. Progress in candidate phyla radiation (CPR) research[J]. *Acta Microbiologica Sinica*, 2020, 60(6): 1284-1303 (in Chinese).
- [10] CASTELLE CJ, MÉHEUST R, JAFFE AL, SEITZ K, GONG XZ, BAKER BJ, BANFIELD JF. Protein family content uncovers lineage relationships and bacterial pathway maintenance mechanisms in DPANN archaea[J]. *Frontiers in Microbiology*, 2021, 12: 660052.
- [11] 蒋建东, 纪彦晗, 王保战. 候选门级辐射类群(candidate phyla radiation)细菌的生理生态与进化[J]. *微生物学杂志*, 2021, 41(6): 1-10.
- JIANG JD, JI YH, WANG BZ. The physiology, ecology and evolution of candidate phyla radiation bacteria[J]. *Journal of Microbiology*, 2021, 41(6): 1-10

- (in Chinese).
- [12] 姜凯, 曹春玲. 微生物暗物质分离培养——以 CPR 和 DPANN 类群为例[J]. 微生物学杂志, 2023, 43(2): 96-105.
- JIANG K, CAO CL. Isolation and cultivation of microbial dark matter—take CPR and DPANN group as examples[J]. Journal of Microbiology, 2023, 43(2): 96-105 (in Chinese).
- [13] BAKER BJ, HUGENHOLTZ P, DAWSON SC, BANFIELD JF. Extremely acidophilic protists from acid mine drainage host *Rickettsiales*-lineage endosymbionts that have intervening sequences in their 16S rRNA genes[J]. Applied and Environmental Microbiology, 2003, 69(9): 5512-5518.
- [14] BAKER BJ, TYSON GW, WEBB RI, FLANAGAN J, HUGENHOLTZ P, ALLEN EE, BANFIELD JF. Lineages of acidophilic archaea revealed by community genomic analysis[J]. Science, 2006, 314(5807): 1933-1935.
- [15] TANAKA N, MEINEKE B, SHUMAN S. RtcB, a novel RNA ligase, can catalyze tRNA splicing and HAC1 mRNA splicing *in vivo*[J]. Journal of Biological Chemistry, 2011, 286(35): 30253-30257.
- [16] SALMAN V, AMANN R, SHUB DA, SCHULZ-VOGT HN. Multiple self-splicing introns in the 16S rRNA genes of giant sulfur bacteria[J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(11): 4203-4208.
- [17] CASTELLE CJ, WRIGHTON KC, THOMAS BC, HUG LA, BROWN CT, WILKINS MJ, FRISCHKORN KR, TRINGE SG, SINGH A, MARKILLIE LM, TAYLOR RC, WILLIAMS KH, BANFIELD JF. Genomic expansion of domain Archaea highlights roles for organisms from new phyla in anaerobic carbon cycling[J]. Current Biology, 2015, 25(6): 690-701.
- [18] LUEF B, FRISCHKORN KR, WRIGHTON KC, HOLMAN HY N, BIRARDA G, THOMAS BC, SINGH A, WILLIAMS KH, SIEGERIST CE, TRINGE SG, DOWNING KH, COMOLLI LR, BANFIELD JF. Diverse uncultivated ultra-small bacterial cells in groundwater[J]. Nature Communications, 2015, 6: 6372.
- [19] BAKER BJ, COMOLLI LR, DICK GJ, HAUSER LJ, HYATT D, DILL BD, LAND ML, VERBERKMOES NC, HETTICH RL, BANFIELD JF. Enigmatic, ultrasmall, uncultivated archaea[J]. Proceedings of the National Academy of Sciences of the United States of America, 2010, 107(19): 8806-8811.
- [20] TRINGE SG, RUBIN EM. Metagenomics: DNA sequencing of environmental samples[J]. Nature Reviews Genetics, 2005, 6(11): 805-814.
- [21] KARLSSON FH, TREMAROLI V, NOOKAEW I, BERGSTRÖM G, BEHRE CJ, FAGERBERG B, NIELSEN J, BÄCKHED F. Gut metagenome in European women with normal, impaired and diabetic glucose control[J]. Nature, 2013, 498(7452): 99-103.
- [22] NIELSEN HB, ALMEIDA M, JUNCKER AS, RASMUSSEN S, LI JH, SUNAGAWA S, PLICHTA DR, GAUTIER L, PEDERSEN AG, LE CHATELIER E, PELLETIER E, BONDE I, NIELSEN T, MANICHANH C, ARUMUGAM M, BATTO JM, QUINTANILHA dos SANTOS MB, BLOM N, BORRUEL N, BURGDORF KS, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes[J]. Nature Biotechnology, 2014, 32(8): 822-828.
- [23] LI DH, LIU CM, LUO RB, SADAKANE K, LAM TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph[J]. Bioinformatics, 2015, 31(10): 1674-1676.
- [24] QIN JJ, LI RQ, RAES J, ARUMUGAM M, BURGDORF KS, MANICHANH C, NIELSEN T, PONS N, LEVENEZ F, YAMADA T, MENDE DR, LI JH, XU JM, LI SC, LI DF, CAO JJ, WANG B, LIANG HQ, ZHENG HS, XIE YL, et al. A human gut microbial gene catalogue established by metagenomic sequencing[J]. Nature, 2010, 464(7285): 59-65.
- [25] MENDE DR, WALLER AS, SUNAGAWA S, JÄRVELIN AI, CHAN MM, ARUMUGAM M, RAES J, BORK P. Assessment of metagenomic assembly using simulated next generation sequencing data[J]. PLoS One, 2012, 7(2): e31386.
- [26] FU LM, NIU BF, ZHU ZW, WU ST, LI WZ. CD-HIT: accelerated for clustering the next-generation sequencing data[J]. Bioinformatics, 2012, 28(23): 3150-3152.
- [27] SUNAGAWA S, COELHO LP, CHAFFRON S, KULTIMA JR, LABADIE K, SALAZAR G, DJAHANSCHIRI B, ZELLER G, MENDE DR, ALBERTI A, CORNEJO-CASTILLO FM, COSTEA PI, CRUAUD C, D'OVIDIO F, ENGELEN S, FERRERA I, GASOL JM, GUIDI L, HILDEBRAND F, KOKOSZKA F, et al. Structure and function of the global ocean microbiome[J]. Science, 2015, 348(6237):

- 1261359.
- [28] BUCHFINK B, XIE C, HUSON DH. Fast and sensitive protein alignment using DIAMOND[J]. *Nature Methods*, 2015, 12(1): 59-60.
- [29] HUSON DH, MITRA S, RUSCHEWEYH HJ, WEBER N, SCHUSTER SC. Integrative analysis of environmental sequences using MEGAN4[J]. *Genome Research*, 2011, 21(9): 1552-1560.
- [30] BERG J, BRANDT KK, AL-SOUD WA, HOLM PE, HANSEN LH, SØRENSEN SJ, NYBROE O. Selection for Cu-tolerant bacterial communities with altered composition, but unaltered richness, via long-term Cu exposure[J]. *Applied and Environmental Microbiology*, 2012, 78(20): 7438-7446.
- [31] MAGOČ T, SALZBERG SL. FLASH: fast length adjustment of short reads to improve genome assemblies[J]. *Bioinformatics*, 2011, 27(21): 2957-2963.
- [32] BOKULICH NA, SUBRAMANIAN S, FAITH JJ, GEVERS D, GORDON JI, KNIGHT R, MILLS DA, CAPORASO JG. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing[J]. *Nature Methods*, 2013, 10(1): 57-59.
- [33] EDGAR RC, HAAS BJ, CLEMENTE JC, QUINCE C, KNIGHT R. UCHIME improves sensitivity and speed of chimera detection[J]. *Bioinformatics*, 2011, 27(16): 2194-2200.
- [34] WANG YY, GUO H, GAO XG, WANG JH. The intratumor microbiota signatures associate with subtype, tumor stage, and survival status of esophageal carcinoma[J]. *Frontiers in Oncology*, 2021, 11: 754788.
- [35] OREN A, GARRITY GM. Valid publication of the names of forty-two phyla of prokaryotes[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2021, 71(10): 005056.
- [36] MURUGKAR PP, COLLINS AJ, CHEN T, DEWHIRST FE. Isolation and cultivation of candidate phyla radiation *Saccharibacteria* (TM7) bacteria in coculture with bacterial hosts[J]. *Journal of Oral Microbiology*, 2020, 12(1): 1814666.
- [37] WURCH L, GIANNONE RJ, BELISLE BS, SWIFT C, UTTURKAR S, HETTICH RL, REYSENBACH AL, PODAR M. Genomics-informed isolation and characterization of a symbiotic *Nanoarchaeota* system from a terrestrial geothermal environment[J]. *Nature Communications*, 2016, 7: 12115.
- [38] HAMM JN, ERDMANN S, ELOE-FADROSH EA, ANGELONI A, ZHONG L, BROWNLEE C, WILLIAMS TJ, BARTON K, CARSWELL S, SMITH MA, BRAZENDALE S, HANCOCK AM, ALLEN MA, RAFTERY MJ, CAVICCHIOLI R. Unexpected host dependency of Antarctic *Nanoarchaeota*[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2019, 116(29): 14661-14670.
- [39] PROBST AJ, LADD B, JARETT JK, GELLER-MCGRATH DE, SIEBER CMK, EMERSON JB, ANANTHARAMAN K, THOMAS BC, MALMSTROM RR, STIEGLMEIER M, KLINGL A, WOYKE T, RYAN MC, BANFIELD JF. Differential depth distribution of microbial function and putative symbionts through sediment-hosted aquifers in the deep terrestrial subsurface[J]. *Nature Microbiology*, 2018, 3(3): 328-336.
- [40] CALLAHAN BJ, WONG J, HEINER C, OH S, THERIOT CM, GULATI AS, MCGILL SK, DOUGHERTY MK. High-throughput amplicon sequencing of the full-length 16S rRNA gene with single-nucleotide resolution[J]. *Nucleic Acids Research*, 2019, 47(18): e103.
- [41] AMIR A, McDONALD D, NAVAS-MOLINA JA, KOPYLOVA E, MORTON JT, XU ZZ, KIGHTLEY EP, THOMPSON LR, HYDE ER, GONZALEZ A, KNIGHT R. Deblur rapidly resolves single-nucleotide community sequence patterns[J]. *mSystems*, 2017, 2(2): e00191-e00116.
- [42] CROSS KL, CAMPBELL JH, BALACHANDRAN M, CAMPBELL AG, COOPER CJ, GRIFFEN A, HEATON M, JOSHI S, KLINGEMAN D, LEYS E, YANG Z, PARKS JM, PODAR M. Targeted isolation and cultivation of uncultivated bacteria by reverse genomics[J]. *Nature Biotechnology*, 2019, 37(11): 1314-1321.
- [43] BOR B, COLLINS AJ, MURUGKAR PP, BALASUBRAMANIAN S, TO TT, HENDRICKSON EL, BEDREE JK, BIDLACK FB, JOHNSTON CD, SHI W, McLEAN JS, HE X, DEWHIRST FE. Insights obtained by culturing *Saccharibacteria* with their bacterial hosts[J]. *Journal of Dental Research*, 2020, 99(6): 685-694.
- [44] NIE J, UTTER DR, KERNS KA, LAMONT EI, HENDRICKSON EL, LIU J, WU TX, HE XS, McLEAN J, BOR B. Strain-level variation and diverse host bacterial responses in episymbiotic *Saccharibacteria*[J]. *mSystems*, 2022, 7(2): e0148821.
- [45] YAKIMOV MM, MERKEL AY, GAISIN VA,



- PILHOFER M, MESSINA E, HALLSWORTH JE, KLYUKINA AA, TIKHONOVA EN, GORLENKO VM. Cultivation of a vampire: 'Candidatus *Absconditicoccus praedator*'[J]. *Environmental Microbiology*, 2022, 24(1): 30-49.
- [46] KURODA K, YAMAMOTO K, NAKAI R, HIRAKATA Y, KUBOTA K, NOBU MK, NARIHIRO T. Symbiosis between *Candidatus* Patescibacteria and archaea discovered in wastewater-treating bioreactors[J]. *mBio*, 2022, 13(5): e0171122.
- [47] KURODA K, KUBOTA K, KAGEMASA S, NAKAI R, HIRAKATA Y, YAMAMOTO K, NOBU MK, NARIHIRO T. Novel cross-domain symbiosis between *Candidatus* Patescibacteria and hydrogenotrophic methanogenic archaea *Methanospirillum* discovered in a methanogenic ecosystem[J]. *Microbes and Environments*, 2022, 37(4): ME22063.
- [48] HUBER H, HOHN MJ, RACHEL R, FUCHS T, WIMMER VC, STETTER KO. A new phylum of archaea represented by a nanosized hyperthermophilic symbiont[J]. *Nature*, 2002, 417(6884): 63-67.
- [49] la CONO V, MESSINA E, ROHDE M, ARCADI E, CIORDIA S, CRISAFI F, DENARO R, FERRER M, GIULIANO L, GOLYSHIN PN, GOLYSHINA OV, HALLSWORTH JE, LA SPADA G, MENA MC, MERKEL AY, SHEVCHENKO MA, SMEDILE F, SOROKIN DY, TOSHCHAKOV SV, YAKIMOV MM. Symbiosis between nanohaloarchaeon and haloarchaeon is based on utilization of different polysaccharides[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2020, 117(33): 20223-20234.
- [50] SAKAI HD, NUR N, KATO S, YUKI M, SHIMIZU M, ITOH T, OHKUMA M, SUWANTO A, KUROSAWA N. Insight into the symbiotic lifestyle of DPANN archaea revealed by cultivation and genome analyses[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2022, 119(3): e2115449119.