



请通过网上投稿系统投稿 网址:http://www.aes.org.cn

基于16S rDNA测序的巢湖流域水体粪便污染溯源

祁钊,赵相龙,桑金慧,何振杰,傅丹丹,岳振宇,宋祥军

引用本文:

祁钊, 赵相龙, 桑金慧, 何振杰, 傅丹丹, 岳振宇, 宋祥军. 基于16S rDNA测序的巢湖流域水体粪便污染溯源[J]. 农业环境科学 学报, 2023, 42(5): 1128–1138.

在线阅读 View online: https://doi.org/10.11654/jaes.2022-0623

您可能感兴趣的其他文章

Articles you may be interested in

基于宏基因组方法分析养猪发酵床微生物组季节性变化

陈倩倩, 刘波, 王阶平, 朱育菁, 张海峰 农业环境科学学报. 2018, 37(6): 1240-1247 https://doi.org/10.11654/jaes.2017-1330

南京地区斑点叉尾养殖池塘水体微生物群落结构研究

钟立强, 王明华, 张世勇, 姜虎成, 陈校辉, 朱广伟, 边文翼 农业环境科学学报. 2020, 39(7): 1594-1604 https://doi.org/10.11654/jaes.2020-0157

生物菌剂对土壤微生物群落结构和功能的影响

沙月霞,黄泽阳,李云翔,赵沛 农业环境科学学报.2022,41(12):2752-2762 https://doi.org/10.11654/jaes.2022-1042

洪湖养殖区水环境中微生物的耐药性及其群落功能多样性研究

关川, 童蕾, 秦丽婷, 刘慧 农业环境科学学报. 2018, 37(8): 1748-1757 https://doi.org/10.11654/jaes.2017-1460

生态循环水养殖池塘抗生素抗性基因污染特征研究

曲疆奇,张清靖,吴彦飞,俞文钰,赵萌,朱华 农业环境科学学报. 2023, 42(3): 641-651 https://doi.org/10.11654/jaes.2022-0492



关注微信公众号,获得更多资讯信息

祁钊,赵相龙,桑金慧,等.基于16S rDNA测序的巢湖流域水体粪便污染溯源[J].农业环境科学学报,2023,42(5):1128-1138. QI Z, ZHAO X L, SANG J H, et al. Tracking fecal contamination in the Chaohu Lake basin based on 16S rDNA sequencing[J]. *Journal of Agro-Environment Science*, 2023, 42(5): 1128-1138.

基于16S rDNA测序的巢湖流域水体粪便污染溯源

祁钊1,赵相龙2,桑金慧2,何振杰2,傅丹丹2,岳振宇1,宋祥军2*

(1.安徽农业大学信息与计算机学院,合肥 230036; 2.安徽省动物性食品质量与生物安全工程实验室,合肥 230036)

摘 要:排入地表水中的动物粪便会带来一系列生态与公共卫生问题,快速准确鉴别污染来源对于源头控制与污染治理具有重要意义。基于细菌群落的微生物溯源技术(Community-based microbial source tracking)以及高通量 DNA 测序技术(Next-generation sequencing, NGS),对污染源中微生物和环境样本中的微生物群落进行比较分析,进而对水体中粪便污染来源进行预测。本文利用 16S rDNA 测序,系统分析与比较了巢湖流域水体、沉积物样本与广泛的潜在污染源(包括村庄与养猪场污水,野生水鸟粪便,人类与家禽、家畜粪便)的细菌群落组成,同时,利用基于机器学习的溯源软件 FEAST 与 Sourcetracker,对水体、沉积物样本的潜在污染源进行了预测。结果表明:水体与沉积物样本的微生物多样性显著高于粪便样本,其中巢湖水体与河流沉积物样本的潜在污染源进行了预测。结果表明:水体与沉积物样本的微生物多样性显著高于粪便样本,其中巢湖水体与河流沉积物样本具有最高的物种多样性,同时样本中也存在大量未分类物种。变形菌门(Proteobacteria)、放线菌门(Actinobacteria),拟杆菌门(Bacteroidetes)广泛分布于所有样本中。溯源分析结果表明,村庄排污口与污水处理厂排污口样本是河水样本最主要的污染。 关键词:16S rDNA测序;微生物群落;水体污染;微生物溯源

中图分类号:X522 文献标志码:A 文章编号:1672-2043(2023)05-1128-11 doi:10.11654/jaes.2022-0623

Tracking fecal contamination in the Chaohu Lake basin based on 16S rDNA sequencing

QI Zhao¹, ZHAO Xianglong², SANG Jinhui², HE Zhenjie², FU Dandan², YUE Zhenyu¹, SONG Xiangjun^{2*}

(1. School of Information and Computer, Anhui Agricultural University, Hefei 230036, China; 2. Anhui Province Engineering Laboratory for Animal Food Quality and Bio-safety, College of Animal Science and Technology, Anhui Agricultural University, Hefei 230036, China) **Abstract**: The discharge of animal waste into rivers can cause a series of ecological and public health problems, thus the rapid and accurate identification of pollution sources is of great importance for source control and pollution management. Combining next-generation sequencing(NGS) and community-based microbial source tracking(MST), we are able to compare the microbial community composition in contamination sources and environmental samples to predict the source of contamination. Using 16S rDNA sequencing, we analyzed the bacterial community composition of water bodies, sediments and potential pollution sources (including village sewage outlets, pig farm effluent, and wild bird, human, poultry, and livestock feces) in the Chaohu Lake basin and analyzed the potential pollution sources of water bodies and sediment samples using machine learning-based traceability software, FEAST and Sourcetracker. The results showed that the microbial diversity of water and sediment samples was significantly higher than that of fecal samples. Chaohu Lake water and river sediment samples exhibited the highest microbial diversity, as well as the presence of a large number of unclassified species. Proteobacteria, Actinobacteria, and Bacteroidetes were widely distributed in all samples. The results of source analysis showed that village

收稿日期:2022-06-22 录用日期:2023-01-07

作者简介:祁钊(1990—),男,安徽濉溪人,博士,实验师,从事智慧农业与溯源检测研究。E-mail:qizhao1050@ahau.edu.cn

^{*}通信作者:宋祥军 E-mail:sxj@ahau.edu.cn

基金项目:国家自然科学基金青年科学基金项目(32202891);安徽高校协同创新项目(GXXT-2019-035)

Project supported: The National Natural Science Foundation of China(32202891); The University Synergy Innovation Program of Anhui Province(GXXT-2019-035)

sewage outlets and wastewater treatment plants were the most important sources of contamination in river water samples. While sediment and lake water samples were potentially contaminated by sewage and wild waterfowl feces, no contamination from human and chicken feces was detected in all samples.

Keywords: 16S rDNA sequencing; microbial community; water pollution; microbial source tracking

环境中的粪便污染是全球范围内日益严重的问 题,尤其是在河流和溪流中凹。流域的粪便污染可能 存在多种来源,包括污水基础设施等点源排污,受上 游牲畜、宠物和野生动物排泄物污染的径流四等。粪 便污染带来的高营养和微生物负荷不仅会影响生态 健康,而且粪便病原体通过水媒的传播还会进一步影 响人类健康。对主要水媒传播病原体的检测是表征 人类健康风险的最直接方法,但由于环境中病原种类 繁多(包括病毒、细菌和原生动物),对其进行检测不 仅会浪费大量时间,还会消耗大量经济成本。为解决 这一难题,监管机构使用粪便指示菌(Fecal indicator bacteria, FIB),如大肠杆菌(Escherichia coli)和粪大肠 菌群(Fecal coliforms)等来评估粪便污染水平^[3],并以 此作为水体微生物风险的替代指标^[4]。然而,传统粪 便指示菌的水质指标检测虽节约了时间和成本,却也 存在诸多不足,如无法提供污染源信息、且无法反映 水体近期的污染情况[5-6]等,近年来许多污染溯源研 究表明,FIB与环境病原相关性不佳[7-8],故而人们开 始探索更加快速、经济的微生物溯源技术。

微生物来源追踪(Microbial source tracking, MST) 技术利用人和动物胃肠道中的微生物设计出具有特 异性的标记物,并通过定量聚合酶链反应(gPCR)等 分子技术检测与宿主相关的微生物标记基因,从而识 别潜在的粪便污染源19-101,但地理差异会显著影响基 于宿主特异性分子标记检测方法的灵敏度和特异性, 这给MST技术的实际应用带来了困难。在过去的20 a中,伴随着下一代测序(Next generation sequencing, NGS)技术的快速发展,研究者对人类、家畜和野生动 物肠道微生物组的理解不断加深凹。使用NGS数据 获取环境及粪便来源的独特的微生物群落图谱的方 法被称为基于细菌群落的 MST (Community-based MST)^[12],主要方法包括基于贝叶斯分类器的Source-Tracker^[13]与基于快速期望最大化算法的微生物源追 踪(FEAST)^[14]。两个MST软件的运行均需要使用者 提供粪便来源文库(Fecal taxon library, FTL),并基于 该文库中不同的宿主菌群组成对待测样本中微生物 的来源进行预测15,进而评估不同污染源对样品微生 物群落的贡献度^[16]。SourceTracker源解析程序通过 对不同宿主粪便以及环境水样 16S rRNA 基因将不同 粪便来源的微生物群落和待分析的水体微生物群落 当作一个整体,基于贝叶斯算法,识别水样中不同宿 主来源微生物所占比例,解析水样中粪便污染来 源^[13]。目前,研究人员已将此方法应用在不同区域以 识别水环境中不同粪便污染来源及其贡献率,例如, Brown等^[15,17]使用 SourceTracker 成功识别出苏必利尔 湖河口粪便污染主要来自污水处理厂排放的废水 (70%)和海鸥粪便(30%),与qPCR方法所得结果一 致。FEAST是一种新兴的计算工具,可用于同时估计 多个潜在来源和各种粪便输入的相对贡献。相同条 件下,FEAST比SourceTracker 表现出更强的稳定性和 更高的运行速度^[14]。但目前应用 FEAST 软件进行河 流水域粪便污染追溯的研究较少,因此其在粪便污染 追溯方面的准确性尚需进一步验证。

在本研究中,研究者提取了所采集样本的DNA 用于16SrDNA测序,旨在通过对比测序结果,探索不 同采样地点[家禽与废水处理厂(WWTP)、河流水体] 及不同样本中(皖江流域地区野生候鸟、人工养殖的 家猪)细菌群落结构的多样性,以及不同物种粪便微 生物组/环境微生物组的差异。此外,将测得的NGS 数据比对到本地潜在病原数据库,用以表征不同生境 样本中的潜在病原组成,为后续污染溯源工作奠定基 础。而后,通过利用不同来源潜在污染物的NGS数 据集构建FTL文库,应用基于SourceTracker与FEAST 两种程序的MST方法对自然水体样本进行污染溯源 研究,进一步表征水体公共安全风险,为巢湖流域水 环境治理提供理论及数据支持。

1 材料与方法

1.1 研究区域概况

巢湖位于安徽省中部,是中国五大淡水湖之一, 同时也是中国富营养化水平最高的湖泊之一^[18]。巢 湖流域水路网络密集,总流入量的80%以上来自双 桥河、派河和南淝河等12条河流。近年来,伴随巢湖 流域人类活动的增加,巢湖受到了严重污染(包括工 业、农业和居民生活污水来源),同时逐渐发展出富营 养化现象^[19]。 1130 <u>1130</u>

1.2 样品采集及处理

1.2.1 水体样本采集

本研究于2021年10月采集水体、沉积物样本共 63个,包括从巢湖流域的杭埠河(HR)、丰乐河(FR)、 派河(XR)各采集的水样(S)3份,沉积物3份,在杭埠 河采样点上游有水禽养殖地点(HRD)额外采集的水 样3份。在巢湖北侧农业面源污染监测站(巢湖烔炀 镇西宋村污水处理站)采集的处理站入水口(D)及出 水口(E)水样各6份,另外还在监测站周边村庄的排 污口(C)采集了6份水样。另在采集粪便样本时,在 淮北昌农收集猪粪便样本(HB)的采样点采集了该养 殖厂的废水处理系统进水(ND)以及出水(NC)样本 各3份。水样用500 mL无菌塑料瓶收集,沉积物样品 使用抓斗取样器在距水面约100 cm深度处采集。样 本于冰上储存,并在6h内运送至实验室进行处理。 水体样本使用0.22 μm聚碳酸酯膜进行过滤,所得滤 膜储存于-80 ℃以供后续分析。

1.2.2 粪便样本采集

本研究共收集粪便样品 229份, 畜禽粪便(猪和 鸡)样品取自皖江流域内规模化养殖场。本研究一共 采集了6个不同地点的猪粪便样本,其中包括从太湖 县收集的猪粪便样本(PA)、金寨县收集的猪粪便样 本(PI)和岳西县收集的猪粪便样本(PL)各5份, 从蚌 埠固镇收集的猪粪便样本(BG)、淮北昌农收集的猪 粪便样本(HB)、宿州褚兰收集的猪粪便样本(SC)、徐 州昌农收集的猪粪便样本(XZ)各4份, 在养殖场采集 的猪粪便样本共31份。鸡粪便样本采集自皖江流域 内规模化养殖场, 共分为5组, 编号为 C、mC、L、LS和 S, 样本数量分别为9、6、15、15份及 15份, 共60份。 另有鹅粪(F)5份。所有粪便采集均使用消毒的 50 mL收集管, 并储存于冰盒, 8h内送至实验室处理。 1.2.3 DNA提取及宏基因组测序

水体、粪便和土壤样本用 DNeasy PowerSoil[®] Pro Kit 试剂 盒提取 DNA,使用通用引物 341F(5'-CCTACGGGNGGCWGCAG-3')和 805R(5'-GAC-TACHVGGGTATCTAATCC-3')PCR 扩增水中细菌的 16S rDNA 基因高变区(V3~V4),PCR产物经2% 琼脂 糖凝胶电泳验证。在整个 DNA 提取过程中,使用超 纯水,以排除假阳性 PCR 结果作为阴性对照的可能 性。得到的扩增子(PCR产物)经纯化后进行浓度检 测,合格后用于测序,扩增子文库的大小和数量分别 在 Agilent2100 生物分析仪和 Illumina 的文库定量试 剂盒上进行评估。

农业环境科学学报 第42卷第5期

1.2.4 样本构成

为增加数据丰富度及可靠性,除了本研究所采样 本外,还使用其他已公布的巢湖水体数据,如Zhang 等[20]采集的18份巢湖流域水体样本:具有农业和生活 污染背景(Agricultural and domestic pollution, ADPR) 的河流包括双桥河(CR1)、拓皋河(CR2、CR12)、鸡裕 河(CR3)、兆河(CR10)和玉溪河(CR11);工业和生活 污染背景(Industrial and domestic pollution, IDPR)的河 流包括南淝河(CR5)、塘西河(CR7)和杭埠河(CR9); 农业污染背景(Agricultural pollution, APR)的河流包 括烔炀河(CR4)、十五里河(CR6)和派河(CR8);还包 括湖中的6个采样点,南淝河口(CL1)、裕溪河(CL6)、 兆河(CL3)以及东湖(CL4、CL5)和西湖(CL2)中心的 样本。该系列样本采用有机玻璃仪器于水下50 cm 深 度人工采集,每个地点采集3个平行水样,混合成1个 水样,下游处理与本研究相同。此外,本研究还选取了 来自菜子湖与升金湖的白额雁(Anser albifrons)粪便样 本30份[21]、升金湖的白头鹤(Grus monacha)粪便样本 16份[22]。以及 Pan 等[23]采集的来自合肥市的 87 份人类 粪便样品。这些样本的下游处理与本研究相同,且测 序均为靶向细菌16SrDNA基因V3~V4区。本研究采 集水体、粪便样本的测序原始数据已上传至NCBI,并 可通过PRJNA783993进行获取。

1.3 数据分析

1.3.1 FTL文库构建

已有研究表明,FTL文库中是否包含具有明确本 地来源的样本对溯源准确度有较大影响,因此本研究 选取巢湖流域北侧农业面源污染监测站(巢湖烔阳镇 西宋村污水处理站)及周边村庄的排污口样本作为 FTL文库中本地污染来源的代表样品。此外,有研究 报道,粪便文库的高组内变异度会对SourceTracker的 溯源结果产生显著影响^[24]。基于这项原则,本研究将 野生水鸟粪便与污水样本按原始样本类型进行了拆 分,野生水鸟粪便拆分为白头鹤(Grus monacha)与小 白额雁(Anser albifrons)粪便,而所有污水样本按处理 前后及污水来源进行了区分,以期最大程度消除粪便 文库的组内变异对预测结果产生的影响。

1.3.2 基于16S rDNA测序的潜在病原菌丰度评估

通过 VFDB 网站(http://www.mgc.ac.cn/VFs/)以 及病理系统资源整合中心(PATRIC)的病原生物数 据,以及其他研究中的潜在病原数据^[25-29],收集了包 括159个属的潜在病原清单,之后根据属名在美国生 物安全协会网站(https://my.absa.org/tiki-index.php?

page=Riskgroups)进行检索,记录该属内生物安全等 级为二级或三级的物种,于LPSN(The List of Prokaryotic names with Standing in Nomenclature) (https:// www.bacterio.net/)网站下载该物种的参考16Sr RNA 序列,如在LPSN网站无法检索到该物种,则从NCBI (http://www.ncbi.nlm.nih.gov/)的GenBank 中检索并下 载对应的16SrDNA序列。最终构建了包括51个属, 444个物种16S rDNA 序列的潜在病原数据库。将聚 类所得 ASV 序列通过 BLASTN 比对到建立的潜在病原 数据库,阈值设置为相似度≥98%、覆盖度>99%,潜在 病原的相对丰度是通过将鉴定为潜在病原的序列对 应的ASV丰度值与样本总ASV丰度的比值来确定。 1.3.3 基于机器学习的微生物污染溯源解析

将不同类型水体或沉积物(即目标样本)设为 Sink,微生物污染源或来源的样品(即构建的FTL文 库样本)为Source。分别使用SourceTracker2以及 FEAST 计算来自不同(Source)来源(人类及动物粪 便、猪场废水处理系统进水与出水、污水处理厂的进 水和出水、村庄污水口)的微生物群落对汇(Sink)环 境(即巢湖流域的水体、河流沉积物)的潜在贡献。 SourceTracker2与FEAST两种方法的区别在于他们是 基于不同的算法来探究目标样本(Sink)中微生物污 染源或进行污染来源(Source)的分析。根据 Source 样本和 Sink 样本的群落结构分布, 预测 Sink 样本中 来源于各Source样本的组成比例。SourceTracker2与 FEAST的分析使用默认参数进行。每个溯源软件进 行5次独立运行以计算每个潜在源的平均贡献度及 其标准偏差。然后通过相对标准偏差(Relative standard deviation, RSD, D)计算平均贡献度与标准偏差 之间的比率,见公式(1)。

$$D = \frac{S}{\bar{x}} \times 100\% = \frac{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2}}{\frac{n-1}{\bar{x}}} \times 100\%$$
(1)

RSD 可以评估多个模型运行的精密程度^[30]。式 中:S为标准偏差(也可以表示为SD);n为重复次数, \bar{x} 表示平均值。RSD较高(≥100%)表明预测结果的可 靠性低。

Wilcoxon 秩和检验(Wilcoxon rank sum test)用于 推断两个独立样本所来自的两个总体分布位置是否 有差别,通常在数据不是正态分布时使用。通过Wilcoxon 秋和检验可以得到一个正态随机变量 Z,再用 软件或查正态分布表得到对应的P值。如果P值较 小(比如小于或等于给定的显著性水平)则可以拒绝

零假设。如果P值较大则没有充分的证据来拒绝零 假设,但不意味着接受零假设。即在数据不是正态分 布时,可以使用Wilcoxon 秩和检验验证样本之间的差 异是否显著。

1.3.4 测序数据处理

数据分析参考扩增子分析流程 EasyAmplicon (https://github.com / YongxinLiu / Easy Amplicon) 完 成^[21]。使用 USEARCH^[22]合并双端序列,去除 barcode 和引物序列。通过 unoise3 去噪获得单碱基精度 ASV,基于SILVA数据库去除嵌合体序列。根据SIL-VA分类器(version2.2, http://sourceforge.net/projects/ rdp-classifier/)(置信阈值0.7)获取每个ASV对应的 物种分类信息以及代表序列。设置测序深度阈值为 30 000,使用 vegan 包进行等量重抽样,通过 R(4.0.3) 包 amplicon(1.11.1)计算 Shannon、Simpson 及 Chao1 指 数。通过R包vegan(2.5.4)计算样本Bray-Curtis距离 以进行 PCoA 分析。其余图例均通过 OmicStudio tools (https://www.omicstudio.cn/tool)以及R语言进行绘制。

结果与分析 2

2.1 测序数据初步分析

291个样本(包括228个粪便样本,63个水体、沉 积物样本)中,共发现10247548条原始序列,经过序 列去冗余,获得了73 844条独特序列,经过 unoise3去 噪及基于 silva 的去嵌合,聚类生成了包括 21 062条 序列的ASV集合,平均每个样本为412个ASVs,最低 244个ASVs,最高444个ASVs。

2.2 Alpha 多样性分析

为了研究数据集样本的微生物多样性,计算了 Shannon、Simpson以及Chao1指数,总体而言,粪便样 本多样性指数水平均低于水体样本。由图1可见,所 有样本中水体沉积物样本具有最高的 ASV 数量与物 种多样性水平,其次是巢湖水体样本(Wilcoxon 秩和 检验,P<0.001),而禽类粪便(包括野生水鸟、鹅与鸡) 具有最低的 ASV 数量与物种多样性水平(Wilcoxon 秩 和检验,P<0.01)。

2.3 细菌群落结构

经过物种注释,所有样本共注释到31个门(图 2),分布最广泛的门包括变形菌门(Proteobacteria)、 放线菌门(Actinobacteria)、拟杆菌门(Bacteroidetes)、 厚壁菌门(Firmicutes)、疣微菌门(Verrucomicrobia), 其中粪便样本以厚壁菌门为主导(平均相对丰度为 55.89%),水体样本则以变形菌门为主导(平均相对

www.ger.org.cn

1132 农业环境科学学报 第42卷第5期 () GR Observed OTUs Shannon 指数 1.00 P log₁₀(丰度 Abundance) 3.0 log10(手度 Abundance) 0.75 2.5 0.50 2.0 0.25 1.5 THE REAL PROPERTY OF THE PARTY The state of the s At Addition THE REAL PROPERTY OF THE PARTY 機制能化 HA State In: 谢性州朝 大学 **光**港^使 街林 樹蓉 大教教師 新 AND N * / 样本Sample , 样本 Sample Chao1 指数 Simpson 指数 0.3 3.5 og10(丰度 Abundance) log₁₀(丰度 Abundance) 3.0 0.2 2.5 0.1 2.0 Alt affer WHI THE HE ~ 选择的 MATHIN 光推開 IN THE REAL PROPERTY OF HA LUTE IN 杨扬度作 横湖市水 劇拳 -Fijk 彩林 脱带 山本 殿楼 AN THE 样本Sample 样本Sample



丰度为44.13%)。除野生水鸟外,粪便样本门水平细 菌集中分布于厚壁菌门、变形菌门与拟杆菌门中,这 3个门平均相对丰度之和占总数的97%以上,而野生 水鸟中除这3种门外,还含有较高相对丰度的放线菌 门(平均13.83%)。水体样本门水平细菌更为多样, 平均相对丰度在1%以上的有13个门。

在属水平上,假单胞菌属(Pseudomonas, 19.81%)、八叠球菌属(Sporosarcina,11.27%)与节杆 菌属(Arthrobacter,9.45%)是野生水鸟中最主要的属, 大肠埃氏菌-志贺氏菌属(Escherichia-Shigella, 29.46%)是鸡粪中最主要的属,乳杆菌属(Lactobacillus,18.9%)与大肠埃氏菌-志贺氏菌属(13.64%)在鹅 粪中广泛分布,而人粪中表现出拟杆菌属(Bacteroides,18.09%)与 普 雷 沃氏 菌 属 _9 (Prevotella_9, 10.74%)的显著富集,猪粪中则以不动杆菌属(Acinetobacter,23.51%)为主。hgcl_clade 在河流水样 (10.23%)与废水样本(14.24%)中表现出明显富集, 不动杆菌属表现出在河水中的显著富集(10.78%), 此外,河流水样中还发现了较高丰度的CL500-29_marine_group(5.8%)与蓝细菌聚球藻属(Synechococcus, 6.96%)。遗憾的是,基于 silva数据库 (v123版本)的注释结果,本研究的河流沉积物与巢 湖水样两组样本中分别有71.51%与79.47%的拼接 后的测试样本未分类到属。

2.4 Beta 多样性分析

基于 Bray-Curtis 距离对整体数据进行了主坐标 分析(图3),将所有数据分为粪便与水体两个组进行 聚类分析。水体样本的beta多样性分析(图3a)表明, 猪场废水与污水处理厂废水样本表现出较大的组内 变异,相比之下巢湖水体样本和沉积物样本则呈现明 显的聚集趋势。粪便样本中,除去各组中少量离群样 本,被分类为同一组的样本(野生水鸟、鸡粪、鹅粪、猪 粪与人粪)各自均表现出显著聚集(图3b),此外,野 生水鸟与猪粪、鹅粪样本表现出明显的聚集,说明其



Figure 2 Composition of microbial community

群落组成存在一定相似性。

2.5 样本中潜在病原的分布

通过将聚类得到的ASV序列比对到自建病原数 据库,对所有样本中潜在病原进行评估,将与数据库 的BLAST(相似度>97%,覆盖度>99%)比对匹配上的 病原菌按属水平聚类(图4)。结果显示,所有样本中 共注释到57个潜在病原属,包括145个潜在病原物 种,其中13个潜在病原属广泛分布于所有样本中,以 Pseudescherichia、肠球菌属(Enterococcus)、链球菌属 (Streptococcus)与肠杆菌属(Enterobacter)为代表(表 1)。在所有样本中,鹅粪、猪场废水与巢湖水样本潜 在病原总平均相对丰度最高,分别为0.051%、0.016% 与0.011%。从潜在病原数量上来看,含潜在病原属 数量最多的为野生水鸟、废水与人粪(分别为48、47、 46个属),值得注意的是,猪场废水独有3种潜在病原 属,分别为Alloprevotella、福赛坦氏菌(Tannerella)与 密螺旋体属(Treponema),而人粪独有一个潜在病原 属,即萨特菌属(Sutterella)。

2.6 基于细菌群落的微生物溯源

将不同类型水体或沉积物设为Sink,构建的FTL



图 3 样本 beta 多样性差异——主坐标分析

Figure 3 Principal coordinates analysis (PCoA) of the microbial community in different groups



Figure 4 Distribution of potential pathogens among all samples (mean relative abundance)

文库样本作为Source,来判别水体/沉积物样本潜在的污染来源。对于FEAST以及SourceTracker两个溯源软件,分别进行5次独立运行以计算溯源结果的RSD。两个溯源软件对于河流水体污染来源的判定结果较为一致(图5)。

2.7 FEAST与SourceTracker对比分析

就两个软件预测结果的稳定性(RSD)来看(图 6),FEAST的运算结果表现出比SourceTracker更低的 RSD值,其样本平均 RDS值多分布于0.25及0.50左 右,而SourceTracker预测值的 RSD计算结果分布在 1.0及以上的偏多。

2.8 微生物溯源结果

FEAST的预测结果表明,沉积物样本的污染来源

主要可被划分为猪场废水出水、猪场原始废水以及白 头鹤粪便,而相较FEAST而言,SourceTracker对于样 本污染贡献程度的判定更趋于保守。但综合二者的 结果来看,FEAST与SourceTracker溯源软件对于河流 水体污染来源的判定结果一致性较高,均将主要的潜 在污染来源归类为村庄排污口以及污水处理厂排污 口样本,FTL文库中小白额雁、人类以及鸡粪几乎在 所有样本中都不能预测出。对巢湖水污染的预测结 果中,两个溯源软件都判定猪场废水出水是主要的污 染来源。

3 讨论

研究表明,畜禽肠道定殖的微生物主要有厚壁菌

2023年5月

表1 所有样本共有的潜在病原的平均相对丰度(按属聚类)

Table 1 Average relative abundance of potential pathogens common to all samples (clustered at the genus level)

病原体属	鸡粪	鹅粪	人粪	野生水鸟	猪粪	猪场废水	污水	河流沉积物	河流水	巢湖水
Pathogen Genus	Chicken	Goose	Human	Wild birds	Pig	Pig wastewater	Waste water	River sediment	River water	Lake Chaohu
Bacteroides	5.841 38E-09	2.939 36E-06	5.108 31E-04	6.231 83E-06	1.980 51E-06	1.926 31E-06	9.699 00E-06	1.579 26E-05	1.576 50E-06	6.408 75E-07
Brevundimonas	1.752 42E-08	3.919 15E-06	4.009 41E-09	2.845 44E-05	2.440 49E-05	3.017 89Е-05	1.953 37E-05	3.398 40E-06	3.905 43E-06	1.089 49E-05
Brucella	5.841 38E-09	8.524 15E-05	1.603 77E-08	5.098 77E-06	8.411 82E-06	1.464 00E-04	7.257 30E-06	2.858 65E-05	1.719 82E-06	1.538 10E-05
Clostridium	2.400 22E-05	1.273 72E-05	1.493 67 E-04	4.402 67E-04	1.894 90E-04	3.358 21E-04	4.205 16E-06	1.139 46E-05	4.120 41E-06	6.408 75E-05
Comamonas	2.330 71E-06	3.135 32E-05	3.728 75E-07	3.216 43E-05	1.765 42E-04	3.403 16E-05	3.141 66E-04	1.259 41E-05	1.837 34E-04	2.371 24E-05
Enterobacter	2.272 30E-06	7.838 30E-05	1.874 80E-05	6.140 45E-04	1.149 97E-06	2.568 42E-06	6.111 05E-05	4.997 65E-06	1.064 14E-05	*0.890%
Enterococcus	4.772 94E-04	**1.532%	7.329 21E-06	3.260 29E-05	2.055 68E-04	9.631 57E-06	1.559 98E-06	7.996 24E-07	6.449 34E-07	6.408 75E-06
${\it Faecalibacterium}$	2.258 86E-05	3.625 21E-05	8.178 20E-04	5.244 97E-06	9.076 25E-05	8.989 47E-06	3.798 21E-06	7.996 24E-07	2.472 25E-06	8.331 38E-06
Paeniclostridium	5.210 51E-06	7.348 4E-05	3.329 82E-05	1.784 93E-04	2.216 89E-05	1.521 79E-04	2.577 36E-06	1.799 15E-06	1.576 50E-06	2.499 41E-05
Pseudescherichia	*0.500%	**2.545%	3.481 81E-04	3.461 31E-05	*0.128%	1.348 42E-05	5.147 93E-05	2.598 78E-06	2.687 22E-06	6.536 93E-05
Pseudomonas	2.394 97E-07	1.959 57E-06	3.207 53E-08	4.440 86E-06	2.082 72E-05	7.641 05E-05	1.199 15E-04	1.869 12E-04	5.421 03E-05	2.948 03E-05
Staphylococcus	1.168 28E-08	*0.127%	3.207 53E-08	2.394 05E-06	2.065 69E-06	5.136 84E-06	3.459 09E-06	1.999 06E-07	2.794 71E-06	3.845 25E-06
Streptococcus	265 701E-04	*0.788%	5.592 33E-05	5.135 32E-06	*0.171%	*0.125%	3.459 09E-06	3.998 12E-07	3.475 48E-06	4.486 13E-06

注:*代表相对丰度大于0.1%,**代表相对丰度大于1%。

Note:* represents relative abundance greater than 0.1%, and ** represents relative abundance greater than 1%.



Figure 5 Microbial source tracking analysis based on NGS dataset

www.aer.org.cn



图6 溯源软件5个独立运行结果的相对标准偏差

Figure 6 Relative standard deviation generated in 5 independent runs of the traceability software

门(Firmicutes)、拟杆菌门(Bacteroidetes)、变形菌门 (Proteobacteria)和放线菌门(Actinobacteria)等,厚壁 菌及其家族成员在畜禽粪便样品中具有较高的相对 丰度。在所有样本中,水体沉积物样本具有最高的 ASV数量与物种多样性水平。沉积物作为一个 营养丰富的栖息地,为微生物的生长提供了良好条 件^[31],例如微藻、浮游植物、硅藻、细菌、浮游动物以促 进、竞争或共生的方式生活在沉积物上,使水体沉积 物发展出丰富的生物多样性^[32]。而粪便微生物凭借 共生关系存活于宿主肠道,其群落组成受到微生物间 互作及宿主的免疫反应等因素影响,并且动物肠道是 一个高度厌氧的环境,对粪便微生物群落施加了特定 的选择压力,导致粪便微生物群落多样性低于自然水 生环境。

在不同类型的水样本中,细菌群落的多样性存在 显著差异,巢湖水样的菌群丰度明显高于其他水样, 这表明与人类活动相关的较高浓度的有机和无机物

可能会降低河流水生细菌群落的物种丰富度,这一结 果与之前的研究结果一致[33]。对于研究采集的具有 农业和生活污染背景的河流水样而言,物种丰富度和 均匀度的降低可能是由于某些物种的高度富集,这些 物种很好地适应了有机或无机废水中的特定条件,并 且能够通过使用各种营养物质来抵抗或承受环境波 动^[34]。值得注意的是,有相关研究表明,水体中的粪 便污染情况可能会受到季节的影响,不同季节水体中 主要的污染源可能会有所差异[35-36]。此外,在对水体 样本进行 beta 多样性差异(主坐标分析)分析时发现, 猪场废水与污水处理厂废水样本表现出较大的组内 变异,这可能是由于"猪场废水"与"污水"组同时包含 了原始废水与处理后废水,而处理过程对细菌群落产 生了较大影响,因此导致上述差异。对巢湖水的污染 预测结果中,两个溯源软件都判定猪场废水出水是主 要的污染来源,但FEAST预测猪场原始废水贡献了 平均2.19%的污染,而在SourceTracker中仅为0.4%。 值得注意的是,CL3样本呈现出与其他巢湖水样不一

致的污染预测结果,两个溯源软件均预测其有(23.0±0.5)%的污染来自白头鹤,Zhang等¹⁰⁹对该采样点的分析指出,该点与其他湖水样本存在明显差异,其含有较高丰度的厚壁菌门(24.13%),而该菌门在其他湖水样本中几乎没有检出。两个溯源软件一致预测存在高水平白头鹤粪便污染的样本还有CR4(SourceTracker-5.61%与FEAST-9.59%),该样本同样含有较高的厚壁菌门细菌(47.77%),高丰度厚壁菌门细菌的存在或许是该预测结果的主要驱动因素。另外,FEAST的预测结果表明,APR与ADPR背景的河流普遍存在潜在的猪粪污染,污染水平从0.205%(CR3)到4.220%(CR8)。值得注意的是,由于本研究所选污水样本的群落组成在一定程度上可能与粪便样本产生交叉,因此溯源软件对村庄污水处理站以及猪场污水样本较高污染贡献度的预测可能存在一定的误差,存在一定的假阳性率。

对于 FEAST 以及 SourceTracker 两个溯源软件来 说,总体上看,两个基于细菌群落 MST 的溯源软件对 于巢湖水体以及河流沉积物的预测结果存在较大差 异,相较FEAST而言,SourceTracker对于FTL样本污 染贡献程度的判定更趋于保守。在FEAST的预测结 果中,沉积物样本的污染来源主要被划分为猪场废水 出水、白头鹤粪便以及猪场原始废水,以FR组样本为 例,FEAST预测其污染源平均有18.50%来自猪场废 水、8.98%来自白头鹤、4.65%来自猪场原始废水,而 SourceTracker对于FR组的预测则仅有5.27%来自猪 场废水出水、0.49%来自白头鹤、0.81%来自猪场原始 废水。在HR以及XR组中,FEAST预测存在3.98%~ 10.30%的猪场废水出水、3.67%~5.39%的白头鹤以 及1.69%~5.40%的猪场原始废水污染,而在Source-Tracker 的 预测结果中,该比例分别降至 0.04%~ 1.54%、0.08%~0.24%与0.18%~0.70%。

综上所述,尽管基于NGS的MST具有众多优势, 但其在研究设计方面仍然存在较高要求,后续研究 中,存在更少ASV交叉、地理联系更加紧密的FTL文 库或许能使溯源预测结果更加精准。

4 结论

(1)不同生境样本中,微生物多样性存在明显差 异。水样本的微生物群落多样性普遍高于人类及动 物粪便样本,其中水体沉积物与湖泊水体样本微生物 多样性最丰富。样本/宿主类型是主导细菌群落差异 的主要因素,水生环境样本中河流沉积物与污水处理 厂细菌群落存在一定相似性,粪便样本中猪粪、野生 水鸟与鹅粪存在一定相似性。

(2)变形菌门、放线菌门、拟杆菌门与厚壁菌门在 粪便与水体样本中广泛存在,但其相对丰度存在差 异。粪便样本以厚壁菌门(平均相对丰度为55.89%) 为主,水体样本则以变形菌门(平均相对丰度为 44.13%)为主。与本地病原数据库的比对结果显示, 以*Pseudescherichia*、肠球菌属、链球菌属与肠杆菌属 为代表的13种潜在病原在水生环境与粪便样本中广 泛分布,这些跨生境样本出现的潜在病原建议在后续 研究与监测中给予重点关注。

(3)溯源分析结果表明,河水样本最主要的污染来 源是村庄排污口与污水处理厂排污口样本;沉积物与 湖水样本则预测出存在猪场排污水与野生水鸟粪便 的污染;所有样本未检测到来自人粪与鸡粪的污染。

(4)以自然水体及河流沉积物作为目标样本,溯源 软件SourceTracker与FEAST对广泛污染源(包括野生 水鸟、家畜、家禽粪便以及生活与养殖业污水)的潜在污 染贡献度的预测结果总体相似,但相对同一污染源而 言,SourceTracker对污染贡献比例的判定偏低。总体 来看,FEAST的预测结果比SourceTracker拥有更低的 RSD值,因此,FEAST模型对潜在污染源的预测结果更 具可信度。在实际调研中,推荐结合NGS与传统污染 评估方法,以准确探明污染来源,指导水质管理。

参考文献:

- REYNOLDS L J, MARTIN N A, SALA-COMORERA L, et al. Identifying sources of saecal contamination in a small urban stream catchment: a multiparametric approach[J]. Front Microbiol, 2021;12:661954.
- [2] CUI Q, HUANG Y, WANG H, et al. Diversity and abundance of bacterial pathogens in urban rivers impacted by domestic sewage[J]. *Environ Pollut*, 2019, 249:24–35.
- [3] 梁红霞, 余志晟, 刘如铟, 等. 分子标记物在禽类粪便污染溯源中的 研究及应用进展[J]. 生态学报, 2021, 41(3):1006-1014. LIANG H X, YU Z S, LIU R Y, et al. Development and application of molecular markers on microbial source tracking for avian[J]. Acta Ecologica Sinica, 2021, 41(3):1006-1014.
- [4] YANG X, CUI H, LIU X, et al. Water pollution characteristics and analysis of Chaohu Lake basin by using different assessment methods [J]. Environ Sci Pollut Res Int, 2020, 27(15):18168-18181.
- [5] WADE T J, CALDERON R L, BRENNER K P, et al. High sensitivity of children to swimming-associated gastrointestinal illness: results using a rapid assay of recreational water quality[J]. *Epidemiology*, 2008, 19(3):375-383.
- [6] NGUYEN K H, SENAY C, YOUNG S, et al. Determination of wild animal sources of fecal indicator bacteria by microbial source tracking(MST) influences regulatory decisions[J]. Water Res, 2018, 144:424–434.

1138 <u>1138</u>

- [7] WU J, LONG S C, DAS D, et al. Are microbial indicators and pathogens correlated? A statistical analysis of 40 years of research[J]. J Water Health, 2011, 9(2):265–278.
- [8] ADAK A, KHAN M R. An insight into gut microbiota and its functionalities[J]. Cell Mol Life Sci, 2019, 76(3):473-493.
- [9] STALEY Z R, BOYD R J, SHUM P, et al. Microbial source tracking using quantitative and digital PCR to identify sources of fecal contamination in stormwater, river water, and beach water in a great lakes area of concern[J]. *Appl Environ Microbiol*, 2018, 84(20):e01634–18.
- [10] HARWOOD V J, STALEY C, BADGLEY B D, et al. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes[J]. *FEMS Microbiol Rev*, 2014, 38(1):1–40.
- [11] UNNO T, STALEY C, BROWM C M, et al. Fecal pollution: new trends and challenges in microbial source tracking using next-generation sequencing[J]. Environ Microbiol, 2018, 20(9):3132-3140.
- [12] DEVANE M L, WEAVER L, SINGH S K, et al. Fecal source tracking methods to elucidate critical sources of pathogens and contaminant microbial transport through New Zealand agricultural watersheds: a review[J]. J Environ Manage, 2018, 222:293–303.
- [13] KNIGHTS D, KUCZYNSKI J, CHARLSON E S, et al. Bayesian community-wide culture-independent microbial source tracking[J]. Nat Methods, 2011, 8(9):761-763.
- [14] SHENHAV L, THOMPSON M, JOSEPH T A, et al. FEAST: fast expectation-maximization for microbial source tracking[J]. Nat Methods, 2019, 16(7):627-632.
- [15] BROWN C M, MATHAI P P, LOESEKANN T, et al. Influence of library composition on SourceTracker predictions for community-based microbial source tracking[J]. *Environ Sci Technol*, 2019, 53(1):60–68.
- [16] STALEY C, KAISER T, LOBOS A, et al. Application of SourceTracker for accurate identification of fecal pollution in recreational freshwater: a double-blinded study[J]. *Environ Sci Technol*, 2018, 52 (7): 4207-4217.
- [17] BROWN C M, STALEY C, WANG P, et al. A high-throughput DNAsequencing approach for determining sources of fecal bacteria in a lake superior estuary[J]. *Environmental Science and Technology*, 2017, 51(15):8263-8271.
- [18] YANG C, YANG P, GENG J, et al. Sediment internal nutrient loading in the most polluted area of a shallow eutrophic lake (Lake Chaohu, China) and its contribution to lake eutrophication[J]. *Environ Pollut*, 2020, 262:114292.
- [19] ZHANG L, FANG W, LI X, et al. Linking bacterial community shifts with changes in the dissolved organic matter pool in a eutrophic lake [J]. Sci Total Environ, 2020, 719(13):73–87.
- [20] ZHANG L, CHENG Y, GAO G, et al. Spatial-temporal variation of bacterial communities in sediments in Lake Chaohu, a large, shallow eutrophic lake in China[J]. International Journal of Environmental Research and Public Health, 2019, 16(20):39-66.
- [21] LIU G, GONG Z Z, LI Q Y. Variations in gut bacterial communities between lesser white-fronted geese wintering at Caizi and Shengjin lakes in China[J]. *Microbiology Open*, 2020, 9(7):e1037.

- [22] ZHAO G H, ZHOU L Z, DONG Y Q, et al. The gut microbiome of hooded cranes (*Grus monacha*) wintering at Shengjin Lake, China[J]. *Microbiology Open*, 2017, 6(3):e00447.
- [23] PAN R B, ZHANG X L, GAO J Q, et al. Analysis of the diversity of intestinal microbiome and its potential value as a biomarker in patients with schizophrenia: a cohort study[J]. *Psychiatry Research*, 2020, 291: 113260.
- [24] 张冰, 吴林蔚, 文湘华. 全国城市污水处理厂中微生物群落的溯源 分析[J]. 环境科学, 2019, 40(8): 3699-3705. ZHANG B, WU L
 W, WEN X H. Potential source environments for microbial communities in wastewater treatment plants (WWTPs) in China[J]. Environmental Science, 2019, 40(8): 3699-3705.
- [25] SHAO K, YAO X, WU Z, et al. The bacterial community composition and its environmental drivers in the rivers around eutrophic Chaohu Lake, China[J]. BMC Microbiol, 2021, 21(1):179.
- [26] DISEGNI A, BRAUN T, BEN S M, et al. Guided protocol for fecal microbial characterization by 16S rRNA-amplicon sequencing[J]. J Vis Exp, 2018, 133:56845.
- [27] ROGNES T, FLOURI T, NICHOLS B, et al. VSEARCH: a versatile open source tool for metagenomics[J]. Peer J, 2016, 4:e2584.
- [28] LI J, CHEN Q, LI H, et al. Impacts of different sources of animal manures on dissemination of human pathogenic bacteria in agricultural soils[J]. *Environ Pollut*, 2020, 266:115399.
- [29] ROGUET A, ESEN Ö C, EREN A M, et al. FORENSIC: an online platform for fecal source identification[J]. mSystems, 2020, 5(2), e00869– 19.
- [30] XU Y, HAN G, ZHANG H, et al. Application of fast expectation-maximization microbial source tracking to discern fecal contamination in rivers exposed to low fecal inputs[J]. *J Microbiol*, 2022, 60(6):594– 601.
- [31] XV Z F, JI J P, SHI C. Water geochemistry of the Chaohu Lake Basin rivers, China: chemical weathering and anthropogenic inputs[J]. Applied Geochemistry, 2011, 26:S379–S383.
- [32] HORNER M C, CARNEY K M, BOHANNAN B J M. An ecological perspective on bacterial biodiversity[J]. *Proceedings Biological Scienc*es, 2004, 271:113-122.
- [33] COLFORD J M, WADE T J, SCHIFF K C. Water quality indicators and the risk of illness at beaches with nonpoint sources of fecal contamination[J] *Epidemiology*, 2007, 18(1):27-35.
- [34] LAPARA T M, BURCH T R, MCANMARA P J, et al. Tertiary-treated municipal wastewater is a significant point source of antibiotic resistance genes into Duluth-Superior Harbor[J]. *Environmental Science* and Technology, 2011, 45(22):9543–9549.
- [35] 张麟杰,王淦淦,张利兰,等.重庆市河水中粪源微生物污染特征及源解析[J].中国环境科学,2019,39(3):1253-1260. ZHANG L
 J, WANG G G, ZHANG L L, et al. Pollution characteristics and source track of fecal microorganism in the rivers across Chongqing City[J]. China Environmental Science, 2019, 39(3):1253-1260.
- [36] OLIVAS Y, FAULKNER B R. Fecal source tracking by antibiotic resistance analysis on a watershed exhibiting low resistance[J]. *Environmental Monitoring and Assessment*, 2008, 139:15–25.