中国科学：地球科学 SCIENTIA SINICA Terrae 2026 年 第 56 卷 第 1 期： 105–116
论 文
《中国科学》杂志社 SCIENCE CHINA PRESS

全球海洋中大洋螺菌目相关病毒的多样性、分布及生态潜能

沈克桦[1†]，张文静[1,2†]，郑凯阳[1†]，申谨豪[1]，任林义[1]，董越[1]，王子岳[1]，刘云丹[1]，邵红兵[1,3]，Andrew MCMINN[1,4]，汪岷[1,3,5,6*]，梁彦韬[1,3*]

1. 中国海洋大学，海洋生物多样性与进化教育部重点实验室，海洋生命学院，海洋生物多样性与进化研究所，深海圈层与地球系统前沿科学中心，海洋碳中和中心，青岛 266003
2. 海南省生态环境监测中心，海口 571126
3. 中国海洋大学-马来西亚登嘉楼大学海洋联合研究中心，青岛 266101
4. Institute for Marine and Antarctic Studies, University of Tasmania, Hobart 7005, Australia
5. 中国海洋大学海德学院，青岛 266101
6. 青岛大学附属医院，青岛 266003

* 通讯作者，E-mail：mingwang@ouc.edu.cn(汪岷)，liangyantao@ouc.edu.cn(梁彦韬)
† 共同第一作者

2025-02-25收稿，2025-10-10修回，2025-10-13接受，2025-12-04网络版发表
崂山实验室项目(LSKJ202203201)、国家自然科学基金项目(42120104006、42176111)、海洋负碳排放国际大科学计划(ONCE)、青岛市市南区科技计划项目(2023-1-012-ZB)和中央高校基础科研基金项目(202172002、201812002、202072001、2024ZZKY、202461036、Andrew McMinn)资助

摘要 大洋螺菌目(Oceanospirillales)细菌在海洋环境中分布广泛，并在碳氢化合物污染的生物修复中扮演着关键角色。然而，与该类群细菌相关的病毒其特征与作用仍不明确。在本研究中，我们从超过1,500万个病毒和细菌的基因组及基因组片段中，鉴定出3,273个与大洋螺菌目相关的病毒重叠群，其中包括783个高质量的大洋螺菌目相关病毒基因组。大多数大洋螺菌目相关病毒为头尾病毒和丝状病毒(分别为549个和17个)。我们共注释出1,226个假定的辅助代谢基因(AMGs)，这些基因涉及39个主要代谢种类。值得注意的是，其中54个假定AMGs与脂肪族烃、芳香烃及其衍生物的降解相关，这突显了这些病毒在碳氢化合物生物降解中的潜在作用。根据系统发育推断和结构相似性网络分析，病毒和宿主所具有的烃及烃衍生物降解相关基因表现出趋同进化。基于生物地理学分析，大多数大洋螺菌目相关病毒被发现于水生环境，其中在温带和热带海洋表层区丰度最高。本研究揭示了大洋螺菌目相关病毒的多样性、分布和生态潜能，并强调了需要将其纳入全球碳氢化合物降解的评估体系。

关键词 大洋螺菌目相关病毒，宏基因组，烃类降解基因，多样性，分布

1 引言

大洋螺菌目(Oceanospirillales)属于γ-变形菌纲，该纲细菌为革兰氏阴性菌。该目细菌呈现出多样化的生理和生态特性。它们通常是嗜盐或耐盐的，并表现为化能异养的生活方式(Kim等，2007)。大洋螺菌目物种在海洋环境中分布广泛，且在石油污染水体中的丰度显著更高(Sabirova等，2006)。这些细菌能够降解多种石油烃类，这使得它们在石油污染物的生物降解以及海洋环境生态平衡的恢复中具有重要意义(Kostka等，2011)。大洋螺菌目在海洋生态系统中占据着独特的生态位；该目中的一些类群在硫循环中扮演着关键角色，特别是在代

中文引用格式：沈克桦, 张文静, 郑凯阳, 申谨豪, 任林义, 董越, 王子岳, 刘云丹, 邵红兵, McMinn A, 汪岷, 梁彦韬. 2026. 全球海洋中大洋螺菌目相关病毒的多样性、分布及生态潜能. 中国科学: 地球科学, 56(1): 105–116, doi: 10.1360/N072025-0061
英文引用格式：Shen K, Zhang W, Zheng K, Shen J, Ren L, Dong Y, Wang Z, Liu Y, Shao H, McMinn A, Wang M, Liang Y. 2026. Diversity, distribution and ecological potential of Oceanospirillales-associated viruses in the global ocean. Science China Earth Sciences, 69(1): 98–108, https://doi.org/10.1007/s11430-025-1722-8

© 2025《中国科学》杂志社

谢二甲基巯基丙酸内盐(DMSP)方面(Delmont等，2015)。此外，某些成员是与极地藻华相关的细菌群落中的优势种群，它们被证实可合成维生素B12，从而促进藻类生长(Joglar等，2021)。再者，大洋螺菌目也是马里亚纳海沟中最丰富的细菌类群之一，在深度超过10,400米的该区域，其烃类降解细菌的比例是地球任何自然环境中观测到的最高值。这些主要来自海洋油杆菌属(*Oleibacter*)、深海弯曲菌属(*Thalassolituus*)和食烷菌属(*Alcanivorax*)的细菌，在维持超深渊带生态系统的稳定性方面发挥着至关重要的作用(Liu等，2019，2024)。

病毒是海洋中最丰富的生物实体，对微生物群落结构和生物地球化学循环产生重大影响(Middelboe和Brussaard，2017)。病毒主要作为微生物群落的调控者，通过感染细菌、古菌和真核微生物等宿主，导致宿主细胞裂解(Liang等，2021；Liu等，2023)。此外，病毒通过向宿主转移遗传物质来促进水平基因转移(Horizontal gene transfer，HGT)，从而增强微生物组的遗传多样性(Irwin等，2022)。病毒还编码辅助代谢基因(Auxiliary metabolic genes，AMGs)，这些基因在感染过程中参与宿主细胞的代谢重编程(Zhou等，2019)。

近年来，宏基因组学与生物信息学分析的快速发展极大地增进了对来自不同环境病毒的了解，并为理解某些重要病毒谱系及其生态潜能提供了庞大的数据集(Moore等，2022；Zheng等，2023a，2023b，2025a)。尽管特殊的大洋螺菌目在过去的几十年里受到了广泛关注，但人们对其病毒的认识仍然有限(Gao等，2022；Liu等，2022；Zhang等，2023)。目前，仅有22个感染大洋螺菌目的病毒分离株被存入NCBI数据库；这些都是经过培养的分离株，并已纳入本研究的比较分析中(Zhang等，2021；Su等，2023)。然而，全球海洋中大洋螺菌目相关病毒的多样性、进化及其生态作用仍然未知。

为增进对大洋螺菌目相关病毒的认识，本研究通过数据挖掘与筛选步骤，从NCBI GenBank、NCBI Genome(Sayers等，2022)及IMG/VR(v.4)数据库(Camargo等，2023)中共鉴定出3,273个与大洋螺菌目相关的病毒重叠群。其中，GenBank数据库包含未组装的病毒重叠群，而Genome数据库则收录了完整测序的病毒基因组。这种整合方法同时涵盖了片段化数据和完整的参考基因组，从而增强了分析的全面性。该数据集包含783个高质量的大洋螺菌目相关病毒基因组，其中包括22个病毒分离株、202个整合的前病毒以及559个未培养的病毒基因组。研究对大洋螺菌目相关病毒的基因组多样性、分布、进化、假定AMGs及HGT进行了探索，首次在全球尺度上提供了对大洋螺菌目相关病毒的全面解析。

2　材料与方法

2.1　大洋螺菌目相关病毒的数据挖掘

从NCBI分类学数据库下载隶属于大洋螺菌目的基因组序列。基于细菌基因组，使用CRASS(v.1.0.1)(Skennerton等，2013)通过间隔序列识别构建了一个CRISPR间隔序列数据库。将CRISPR间隔序列与病毒重叠群进行比对，标准为仅允许一个错配，或多个错配但需满足≥80%的比对一致性。使用BLASTn(v.2.14.0)通过将构建的间隔序列数据库与IMG/VR(v.4)(https://img.jgi.doe.gov/cgi-bin/vr/main.cgi)数据库(包含超过1500万个病毒基因组和UViGs)进行比对来进行同源性搜索。由此获得了3,273个源自宏基因组的病毒重叠群，并将其假定的宿主归属于大洋螺菌目。这些病毒重叠群的组装质量通过CheckV(v.1.0.1)(Nayfach等，2021)进行评估，最终得到559个被标记为“高质量”和“具有直接/反向末端重复序列(DTR/ITR)”(或组装完整度≥90%)的病毒重叠群。通过geNomad(v.1.7.0)(Camargo等，2024)和CheckV(v.1.0.1)识别并筛选出整合的前病毒，最终得到202个被标记为“高质量”的整合前病毒区域。

2.2　基因组相似性聚类网络的构建

从国际病毒分类委员会(ICTV)中检索到的、具有明确分类学地位的大洋螺菌目相关病毒被用作参考数据集。该分析使用vConTACT2(v.0.11.3)(Jang等，2019)进行。首先，使用Diamond Blastp(v.0.9.14)对所有病毒进行全对全比对，参数设置为：e值<1×10^{-10}，序列一致性>50%，序列覆盖度>50%，最大输出结果数为1,000,000条。随后，应用MCL算法(v.14.137)(Buchfink等，2021)根据获得的结果对蛋白质进行聚类。之后，检查按基因组排列的文件，以识别包含大洋螺菌目相关病毒序列的病毒簇(VCs)，并选择这些病毒簇内的其他成员作为参考。利用ntw文件，筛选出与目标序列具有显著权重关系的序列，以识别那些与大洋螺菌目相关病毒相似的序列。随后，将病毒的相似性关系导入Gephi(v.0.10)(Bastian等，2009)进行可视化，其中所有序列IDs被表示为节点，而来自ntw文件的相关性强度则被用作边的权重。

2.3　病毒分类阶元注释

基于平均核苷酸同一性(ANI)，使用VITAP(v.1.7.0)(Zheng等，2025b)对所有大洋螺菌目相关病毒进行分类。具体的分类标准如下：ANI值高于70%的病毒被归为同一病毒属，而ANI值高于95%的病毒则被划分为同一病毒种。

利用蛋白质聚类的结果，构建了共享蛋白质簇百分比的热图。共享蛋白质簇的百分比通过计算两个基因组之间共享蛋白质簇的数量与包含这两个基因组的蛋白质簇总数之间的平均比率来确定。基于共享蛋白质簇百分比构建的矩阵，使用R语言中的“complete”聚类方法生成热图。通过整合这两种方法，完成了对大洋螺菌目相关病毒分类学等级的注释。

2.4 大洋螺菌目相关病毒的基因功能注释

对所有大洋螺菌目相关病毒所编码的开放阅读框(ORFs)使用KEGG直系同源基因数据库(Kanehisa等，2017)和Pfam数据库(Mistry等，2021)进行注释，阈值设定为*e*值<1×10^{-5}。后续使用VirSorter(v.1.0.6)(Roux等，2015)和DRAM-v(v.1.5.0)(Shaffer等，2020)进行注释。被标记为‘-M’或‘-m’并且伴有‘-E’和/或‘-K’标志的基因，若其辅助评分为1、2或3，则被视为潜在的AMGs。随后，使用iPath3(https://pathways.embl.de)(Darzi等，2018)将注释得到的AMGs映射到它们所参与的通路中，并在科水平上对结果进行了统计分析。

2.5 大洋螺菌目相关病毒的相对丰度及分布

利用CoverM(v.0.6.1)(Aroney等，2025)计算了大洋螺菌目相关病毒在全球海洋病毒组(GOV2.0)数据库中的相对丰度。比对标准设定为序列一致性大于95%且序列覆盖度大于75%。GOV2.0数据库共包含154个站位的病毒宏基因组，这些数据被划分为不同的病毒生态区：北极区(ARC)40个样本，南极区(ANT)5个样本，深海层区(BATHY)10个样本，温带和热带海洋表层区(EPI)85个样本，以及温带和热带海洋中层区(MES)14个样本(Ghiglione等，2012)。采用每千个碱基的转录每百万映射读取的转录本(TPM)进行标准化分析，以评估大洋螺菌目相关病毒序列在全球海洋区域的相对丰度和分布模式。

2.6 大洋螺菌目相关病毒的宿主预测

通过CRISPR间隔序列匹配，从IMG/VR(v.4)数据库中鉴定出目标病毒序列。随后，通过将这些序列与IMG/VR(v.4)数据库提供的宿主注释信息进行比对，获得其宿主信息。此外，采用iPHoP(v.1.3.3)(Roux等，2023)验证IMG/VR预测的宿主信息准确性，并选择具有最高置信度得分的预测结果。两种方法所得结果保持一致。

2.7 病毒介导的HGT网络的构建

遵循先前建立的HGT网络构建分析方法(Schulz等，2020)。首先，对NCBI提供的非冗余蛋白质序列(nr)数据库进行过滤，创建一个仅包含细胞生物；具体包括细菌(NCBI分类数据库ID：2)和古菌(NCBI分类数据库ID：2157)的蛋白质数据库。随后，使用Diamond Blastp将所有大洋螺菌目相关病毒蛋白质序列与该细胞生物蛋白质数据库进行比对，参数设置为：*e*值<1×10^{-50}，序列一致性>50%，序列覆盖度>50%，最大输出结果数为1,000,000条。

使用来自nr数据库的病毒蛋白质数据库，将分类限制为病毒(NCBI分类数据库ID：10239)。采用相同参数，对所有大洋螺菌目相关病毒蛋白质序列与该病毒蛋白质数据库执行Diamond Blastp搜索。对于病毒的Blastp结果，仅保留每个大洋螺菌目相关病毒蛋白质中*e*值最低的匹配用于后续分析。

如果一个大洋螺菌目相关病毒蛋白质在细胞生物蛋白质数据库和病毒蛋白质数据库中均有匹配，且其在细胞生物数据库中的*e*值低于在病毒数据库中的*e*值，则该蛋白质被认为与细胞生物亲缘关系更近并可能来源于它们，即发生了一次潜在的HGT。如果一个大洋螺菌目相关病毒蛋白质仅在细胞生物蛋白质数据库中有匹配，这同样被视为一次潜在的HGT。

对所有潜在的HGT，将细胞生物的分类等级追溯至目水平，并对Blastp的bit-scores进行标准化。标准化后的分数被用作权重，以表示潜在HGT事件中病毒与细胞生物之间的关系，细胞生物的出现频率通过节点大小来表示。最后，将结果导入Gephi进行可视化。

2.8 大洋螺菌目相关病毒编码的烃及烃衍生物降解基因(ovHHDDEGs)的系统发育分析

为研究在大洋螺菌目相关病毒中注释到的ovHHDDEGs的系统发育关系，我们构建了一个最大似然系统发育树。选取了具有代表性的ovHHDDEGs基因*hsd*、*dmpD*和*tcpA*。参考宿主序列来自NCBI GenBank数据库中所有归类为Oceanospirillales的宿主基因组。多序列比对使用MUSCLE(v.5.1)(Edgar，2022)的super5模式完成。所得比对结果随后通过trimAL(v.1.4)(Capella-Gutiérrez等，2009)处理，去除在超过10%序列中存在缺口的位置，以消除高度突变位点。接着，利用IQ-tree(v.2.0.3)(Nguyen等，2015)在LG+F+R10模型下构建最大似然系统发育树，并执行3,000次超快速自举迭代。最终系统发育树通过iTol(https://itol.embl.de)进行可视化。

为探究大洋螺菌目相关病毒中烃降解基因与其宿主细菌同源基因之间的系统发育关系，我们使用mmseqs(v.15.6)(Steinegger和Söding，2017)进行同源性搜索，

参数设定为–cov-mode 0 -c 0.5 –min-seq-id 0.5 -e 1e-10。所得序列基于>10%序列相似性阈值进行聚类，并与病毒编码的烃降解基因合并为一个数据集。系统发育树的构建采用与上述相同的参数设置。最终结果通过iTol进行可视化。

2.9　ovHHDDEGs的蛋白结构相似性网络

使用ESMFold(Lin等，2023)对病毒编码的ovHHDDEGs及大洋螺菌目中的同源基因进行蛋白质结构预测。对于每个蛋白，均生成一个PDB文件，并分别构建病毒与大洋螺菌目蛋白的PDB数据库。随后，利用Foldseek(v.8.ef4e960)(van Kempen等，2024)将病毒编码的烃降解基因蛋白结构与大洋螺菌目的同源蛋白结构进行比较。以TM-score作为权重，并通过Gephi对结果进行可视化。

3　结果

3.1　大洋螺菌目相关病毒的鉴定与多样性

在对超过1,500万条病毒与细菌基因组以及UViGs数据进行挖掘后，共预测得到3,273条与大洋螺菌目相关的病毒重叠群。大洋螺菌目相关病毒主要来源于远洋和海水养殖场样本，这与其宿主的生境相一致(图1B)。然而，除了水生环境之外，大洋螺菌目相关病毒也已在人工环境、宿主相关环境以及陆地环境中被检测到。大洋螺菌目相关病毒在多种海洋环境中均表现出广泛分布，包括南北极地区。值得注意的是，相对丰度最高的大洋螺菌目相关病毒样本采集自南极戴维斯站附近海域(图1A)。这些结果表明，大洋螺菌目相关病毒具备适应多样生态位并在不同环境中持续存在的潜力。

共鉴定出783条完整或高质量(完整度≥90%)的大洋螺菌目相关病毒基因，其中包括22条病毒分离株、202条前病毒以及559条UViGs(网络版附图S1)。基于平均核苷酸同一性(ANI≥95%)，这783条高质量大洋螺菌目相关病毒被划分为572个病毒操作分类单元(vOTUs)(Konstantinidis和Tiedje，2005；Brum等，2015；Gregory等，2016)。在这些vOTUs的分类注释中，518个被归为头尾噬菌体，17个为丝状噬菌体，另有1个为无尾二十面体病毒。在科水平上，共有145个大洋螺菌目相关病毒被归入P2-like病毒(P2噬菌体科，*Peduoviridae*)，其次为T7-like病毒(自转录病毒科，*Autographiviridae*，9个)和丝状病毒科(*Inoviridae*)病毒(6个，图2A)。此外，有285个大洋螺菌目相关病毒未与任何参考病毒聚类，而是形成了多个新的病毒类群(图2A)，提示存在丰富的尚未分类的大洋螺菌目相关病毒。

对大洋螺菌目相关病毒的分类采用了基因共享网络分析。包含已确认分类的参考序列以及本研究数据集中序列的分组(以红色椭圆标注)表现出清晰的聚类模式(图2A)。这些分组很可能属于相同的科，主要包括*Peduoviridae*、*Inoviridae*、*Autographiviridae*和微病毒科(*Microviridae*)，这一结果与VITAP的预测一致。此外，仍有大量病毒在科水平上未能归类，进一步表明大洋螺菌目相关病毒含有丰富而尚未探索的多样性。

对大洋螺菌目相关病毒进行的共享蛋白簇(PCs)比例分析显示，规模最大的共享PCs分组归属于*Peduoviridae*，同时也存在相当比例在科水平上未被归类的病毒(图2B)。这一结果再次表明，目前大多数大洋螺菌目相关病毒尚无法被准确分类，提示它们携带大量未知基因并具有高度多样性。

3.2　大洋螺菌目相关病毒与潜在宿主之间的联系

病毒与宿主会随时间共适应并共进化。通过促进跨多个株系或物种的HGT，病毒成为细菌进化的重要驱动力(Markwitz等，2022)。通过构建HGT网络，可以观察大洋螺菌目相关病毒如何从不同宿主获得基因，从而丰富其病毒基因库。本研究发现的572个vOTUs通过HGT分析与43个不同的细菌目建立了关联。在科水平上，共有847条关联连接了3个病毒科与43个细菌谱系，其中包括大洋螺菌目(Oceanospirillales，488条关联)、交替单胞菌目(Alteromonadales，45条关联)以及假单胞菌目(Pseudomonadales，112条关联)。此外，288个大洋螺菌目相关病毒(288条关联)与多样化的细菌目相关联(图3)。这些结果为大洋螺菌目与其他细菌之间的病毒介导推测HGT提供了网络图谱。

3.3　大洋螺菌目相关病毒的基因和潜在AMGs

通过对大洋螺菌目相关病毒的泛基因组分析，共鉴定出33,319个编码已知功能蛋白的病毒基因。其中，1,226个假定的AMGs涉及39个不同的功能类别，归属于8条主要代谢通路。其中，与半胱氨酸和蛋氨酸代谢(175个)、嘧啶代谢(161个)、嘌呤代谢(125个)以及嘧啶代谢(109个)相关的AMGs数量最为丰富。

不同的大洋螺菌目相关病毒在潜在宿主中展现出多样的AMG编码模式。与共生菌科(Endozoicomonadaceae)相关的大洋螺菌目相关病毒编码的AMGs数量最多且最为多样，共977个AMGs，涉及38种不同的功能类别。其中最丰富的AMGs与嘌呤代谢(66个)、苯丙氨酸代谢(47个)以及半胱氨酸和蛋氨酸代谢(37个)相关。此

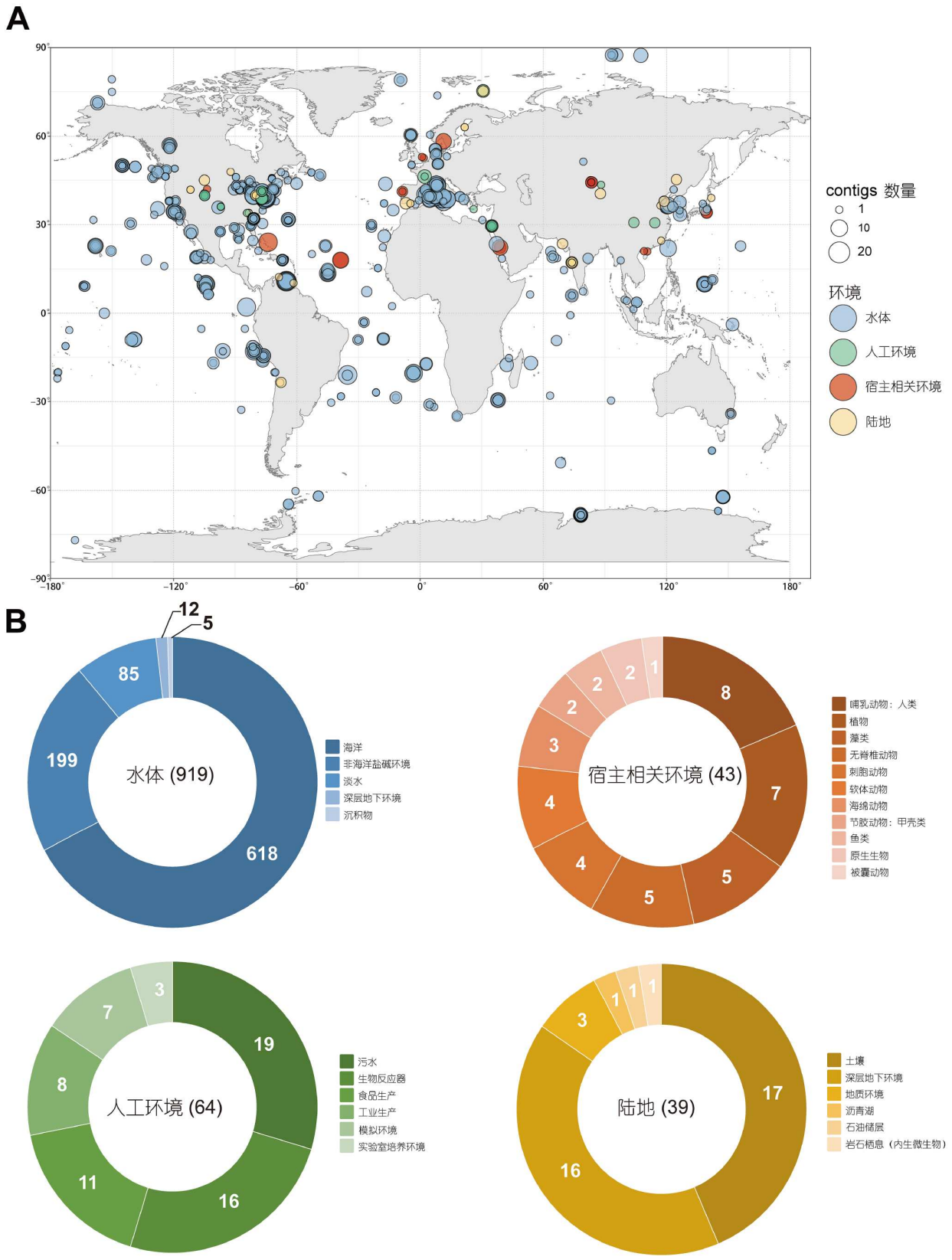

图 1　所有大洋螺菌目相关病毒的地理分布。（A）3,273条重叠群在全球海洋环境中的分布。每个大洋螺菌目相关病毒的检测位置以圆圈表示，其大小与检测到的大洋螺菌目相关病毒数量成正比。不同环境来源的病毒以不同颜色的节点表示。（B）在3,273条重叠群中，展示了1,065条含有生境信息的重叠群，其中四种不同颜色代表四类不同的生境。在饼图中每个生境的扇区内，进一步细分了次级生境来源

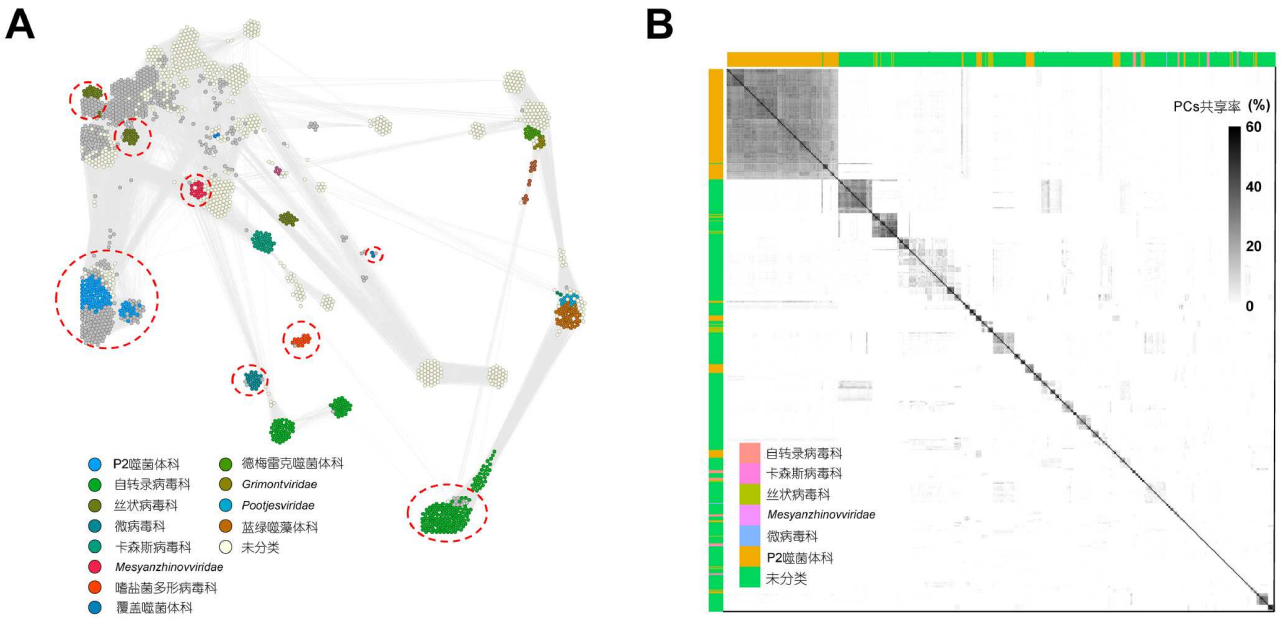

图 2　大洋螺菌目相关病毒的蛋白质和基因组多样性。(A) 基于二分网络分析，展示大洋螺菌目相关病毒与参考病毒基因组之间的关系。不同科的大洋螺菌目相关病毒以不同颜色表示。红色椭圆所标注的群体包含分类地位未知的大洋螺菌目相关病毒，并以灰色表示。(注意：图中标注为“未分类”的白色点代表分类地位未知的参考病毒)。(B) 大洋螺菌目相关病毒之间平均共享蛋白质簇(PCs)的百分比。热图左侧和顶部以不同颜色的方框标注了六个不同的病毒科

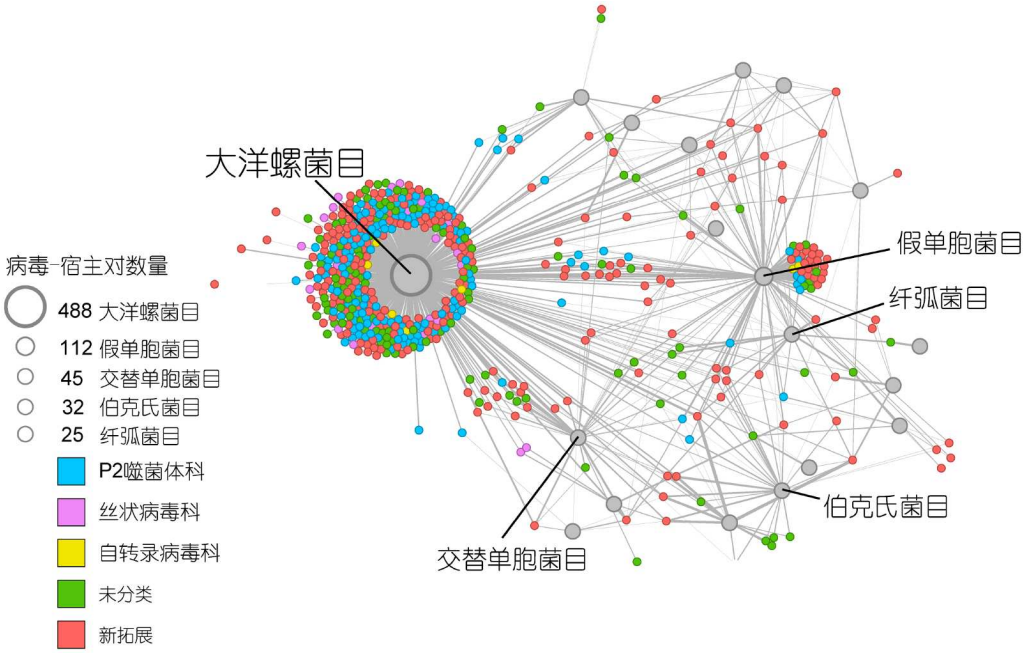

图 3　大洋螺菌目相关病毒与细菌类群的水平基因转移网络。细菌类群以灰色圆圈表示，圆圈大小与病毒-宿主对的数量成正比。病毒以彩色圆圈表示，不同颜色代表不同的病毒科

外，与盐单胞菌科(Halomonadaceae)和烷烃降解菌科(Alcanivoracaceae)相关的大洋螺菌目相关病毒也具有广泛的AMG分布(图4)。

3.4　大洋螺菌目相关病毒及其潜在宿主中烃及其烃衍生物降解基因多样性

值得注意的是，与烃及烃衍生物降解(例如甲苯降

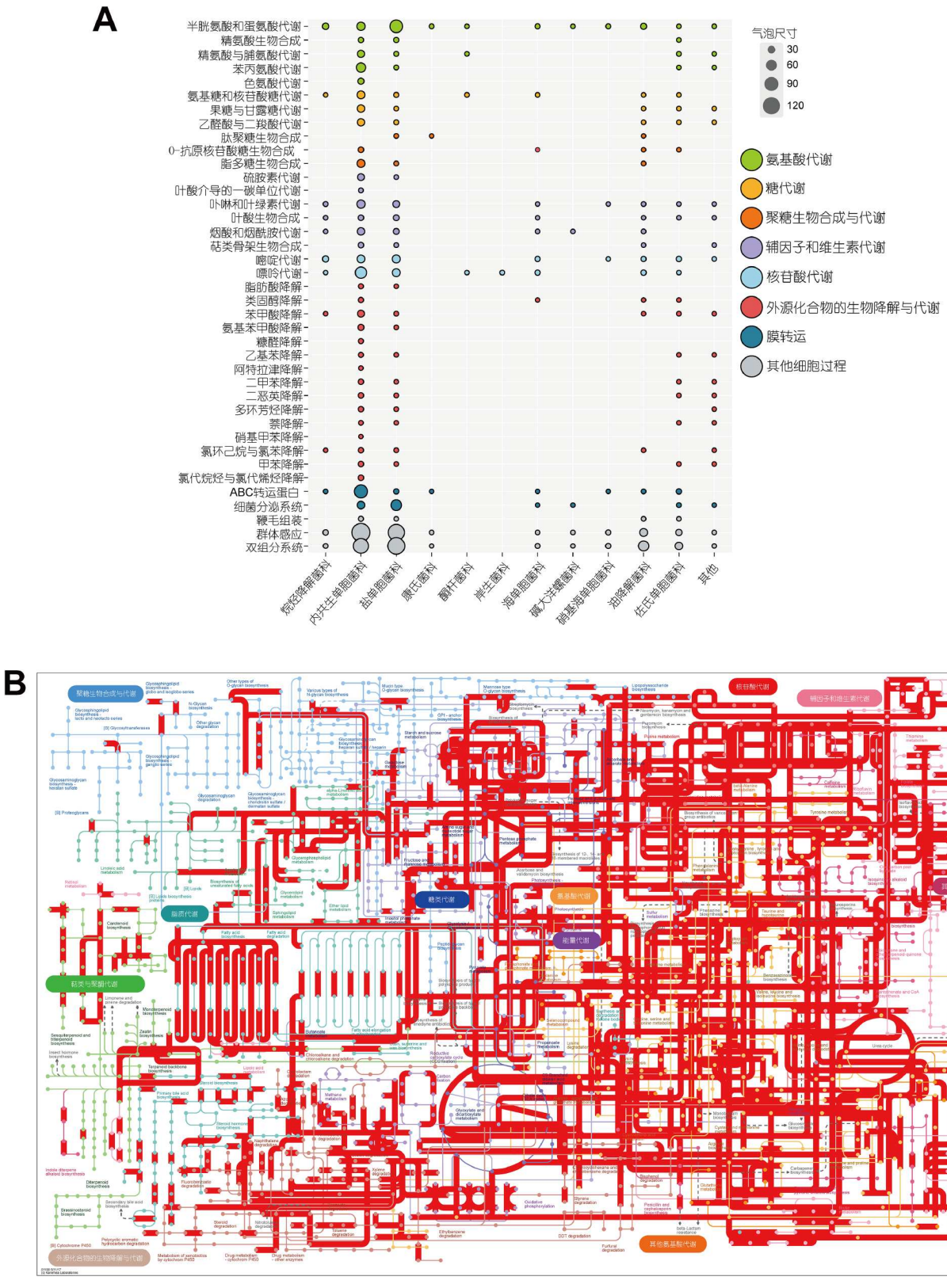

图 4 大洋螺菌目相关病毒的基因功能。（A）大洋螺菌目相关病毒编码的假定辅助代谢基因（Auxiliary metabolic genes，AMGs），包括 1,226个AMGs，涉及39种代谢类型。*x*轴表示感染不同科宿主的大洋螺菌目相关病毒，*y*轴表示不同的代谢类型。热图上方用不同颜色标注了AMGs的不同代谢分类。（B）大洋螺菌目相关病毒编码的假定辅助代谢基因（Auxiliary metabolic genes，AMGs）涉及的代谢通路，其中红色路径表示这些AMGs所参与的相关通路。图中共展示11条代谢通路，每条通路以不同颜色表示

解，16个；二甲苯降解，13个；以及氯环己烷和氯苯降解，12个）相关的大洋螺菌目相关病毒共编码54个AMGs。尤其是与Endozoicomonadaceae和Halomonadaceae相关的大洋螺菌目相关病毒，更倾向于编码与烃及烃衍生物降解相关的AMGs（18个，占3%，图4和图S2A）。在这54个ovHHDDEGs中，包括*badH*（编码2-羟基环己烷羧基辅酶A脱氢酶，6个）、*E3.8.1.2*（编码2-卤代酸脱卤酶，5个）、*pcaD*（编码3-氧代己二酸烯醇内酯酶，4个）、*hsd*（编码3β-羟基-Δ5-类固醇脱氢酶，3个）等（图S2B和图S2C）。其中，*hsd*编码的3β-羟基-Δ5-类固醇脱氢酶参与类固醇降解；*dmpD*（*xylF*）编码的2-羟基粘康酸-半醛水解酶参与苯甲酸降解、二甲苯降解和苯乙烯降解；*tcpA*编码的3,5,6-三氯吡啶-2-醇/2,4,6-三氯苯酚单加氧酶则参与氯环己烷和氯苯降解。这些AMGs被发现与多种烃及烃衍生物的降解相关，包括苯、卤代烃和多环芳烃等。这些烃类均为常见于石油污染海洋环境中的持续性有机污染物。大洋螺菌目相关病毒中存在这些 AMGs，很可能有助于其宿主降解此类化合物，从而帮助大洋螺菌目适应石油污染环境（Markwitz等，2022）。

来自与烃及烃衍生物降解相关的54个AMGs的三个代表性ovHHDDEGs的蛋白质结构相似性和系统发育分析显示，它们与细菌同源基因聚类，TM-score为0.9。代表基因K16045（*hsd*）、K10216（*dmpD*）和K24293（*tcpA*）在系统发育树上与宿主的同源蛋白呈现明显的聚类（图S3）。这表明病毒与潜在宿主的蛋白质结构具有相似的拓扑结构，进一步提示与烃类降解相关的病毒AMGs在功能上可能与其细菌同源基因相似。

这些蛋白及其宿主同源蛋白的最大似然系统发育树显示出趋同的进化关系。在三个蛋白的系统发育树中，K16045与其宿主同源蛋白形成了最大的聚类。K24293与K10216位于同一主干分支上，提示它们具有共同起源，而包含K24293的分支表现出明显的单系群特征。这些结果与蛋白质结构相似性网络的分析结果一致，进一步支持观察到的进化关系。与病毒蛋白K24293发生基因交流的宿主均属于Alcanivoracaceae。相比之下，与病毒蛋白K16045和K10216发生基因交流的主要宿主群则来自Halomonadaceae和Alcanivoracaceae（图5）。

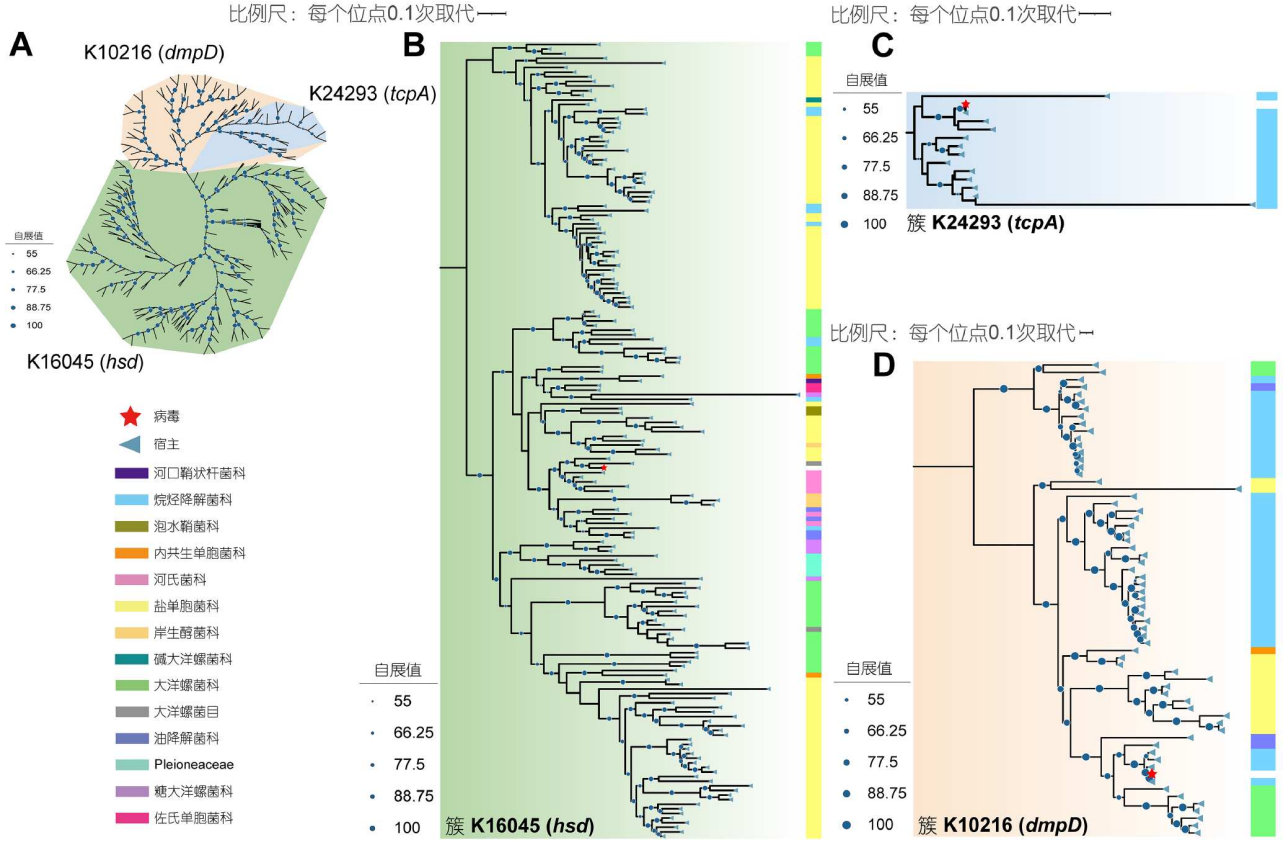

图 5　三个代表性大洋螺菌目相关病毒编码的烃及烃衍生物降解基因（oceanovirus-encoded hydrocarbon and hydrocarbon derivative degradation genes，ovHHDDEGs）与宿主同源基因的最大似然系统发育树。不同宿主科以不同颜色表示。图中显示了系统发育树中包含的分支：（A）所有三个ovHHDDEGs，（B）K16045（*hsd*），（C）K24293（*tcpA*），以及（D）K10216（*dmpD*）

3.5 大洋螺菌目相关病毒的生物地理分布

大洋螺菌目相关病毒分布较为广泛，其空间分布总体上与宿主的分布相一致(图6和图S5)。在GOV2.0数据库中代表五个海洋生态区(viral ecological zones，VEZs)的154个站位的病毒组样本中，计算了572个大洋螺菌目相关病毒vOTUs的相对丰度(Gregory等，2019)。结果显示，大洋螺菌目相关病毒分布广泛，其中EPI丰度最高，其次为ARC，而ANT丰度最低(图6C)。

不同大洋螺菌目相关病毒群体在各VEZs的分布模式存在差异。例如，丝状大洋螺菌目相关病毒(*Inoviridae*)和P2-like大洋螺菌目相关病毒(*Peduoviridae*)在多个VEZs中均有广泛检测，而T7-like大洋螺菌目相关病毒(*Autographiviridae*)在BATHY未被检测到。感染大洋螺菌目的微病毒仅在EPI被检测到(图6B)。在物种水平上，丰度排名前30的大洋螺菌目相关病毒均为UViGs，表明海洋中仍存在大量尚未分离的大洋螺菌目相关病毒(图6A)。

4 讨论

4.1 大洋螺菌目相关病毒的多样性已得到显著扩展

大洋螺菌目物种在海洋环境中分布广泛且丰度较高(Sun等，2020)。目前，仅有22条感染大洋螺菌目的病毒被分离并进行了表征(Zhang等，2021；Su等，2023)，因此培养和未培养的大洋螺菌目相关病毒的基因组多样性及生态分布尚未得到系统研究。本研究基于病毒分离株、整合前病毒及UViGs，结合全球大规模病毒与微生物基因组及宏基因组数据集(Sayers等，2022；Camargo

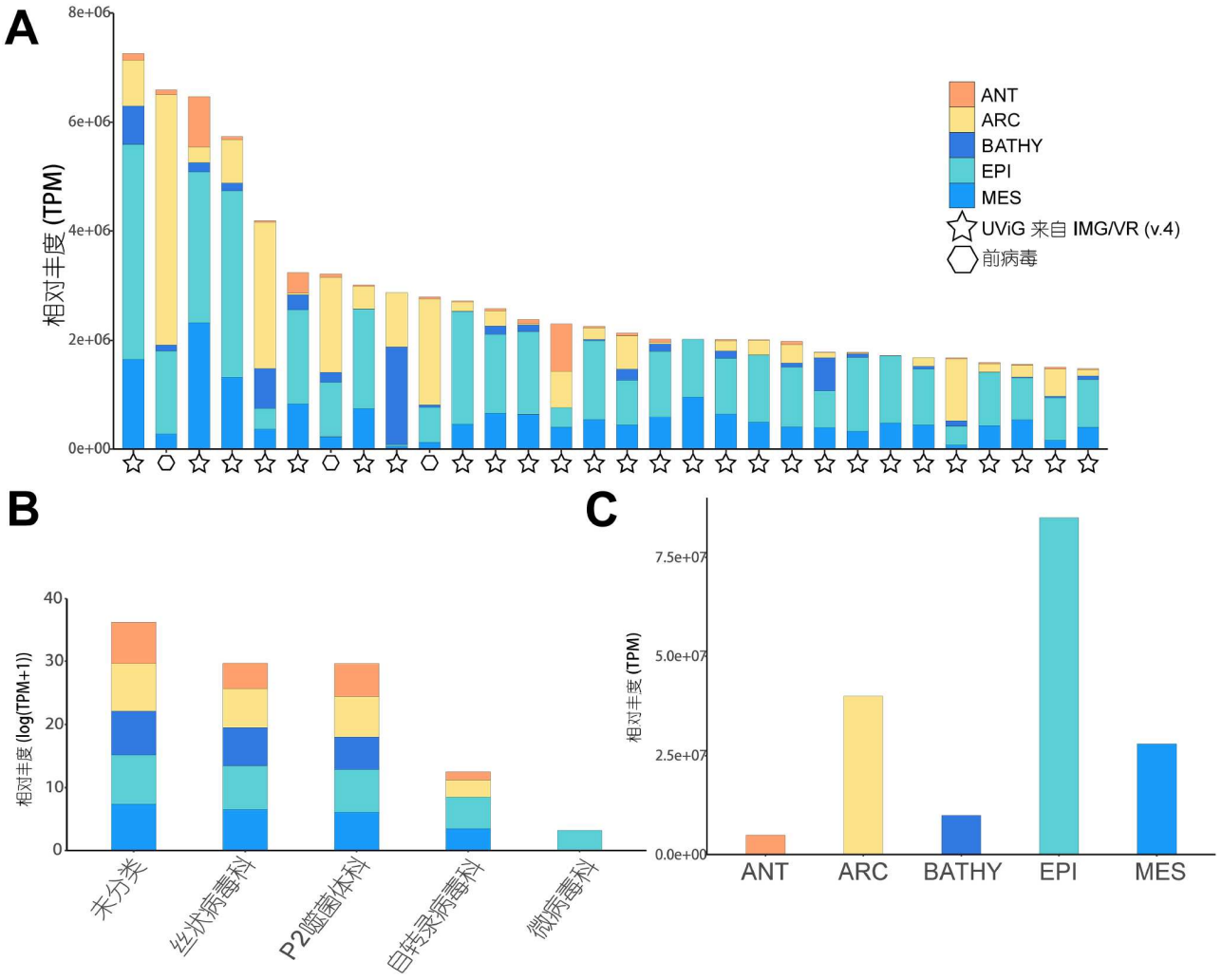

图6 基于GOV2.0数据集中154个站位病毒宏基因组的大洋螺菌目相关病毒相对丰度及分布模式。(A) 在五个病毒生态区(viral ecological zone，VEZs，北极：ARC，南极：ANT，热带及温带表层：EPI，中层：MES，深层：BATHY)中相对丰度最高的前30条大洋螺菌目相关病毒，不同VEZs以不同颜色表示。(B) 四个病毒科中大洋螺菌目相关病毒分布模式的差异。(C) 五个VEZs中大洋螺菌目相关病毒分布模式的差异

等，2023），系统分析了大洋螺菌目相关病毒的基因组、分类及生态特征。

在这22条大洋螺菌目相关病毒分离株中，只有5条在科水平上具有明确分类，其余均未能在该分类级别归类。本研究将大洋螺菌目相关病毒的科水平多样性从原来的2个扩展到5个。除覆盖噬菌体科（*Corticoviridae*）和*Autographiviridae*外，*Peduoviridae*、*Inoviridae*和*Microviridae*也被鉴定为主要的大洋螺菌目相关病毒聚类。此外，这22条分离病毒均与两个宿主科——海单胞菌科（Marinomonadaceae）和Halomonadaceae相关。本研究将大洋螺菌目相关病毒在科水平上的潜在宿主范围从2个扩展至11个（图4）。在基因组多样性方面，大洋螺菌目相关病毒的基因组数量从22条扩展至3,273条，并基于783条高质量基因组划分出572个物种水平的vOTUs（图2）。

4.2　大洋螺菌目相关病毒拥有丰富的AMGs

病毒携带AMGs在病毒感染过程中，对维持或补充宿主细胞的关键代谢功能具有重要作用（Crummett等，2016）。与Endozoicomonadaceae、Halomonadaceae和Alcanivoracaceae相关的病毒表现出最具特异性的AMGs，主要参与半胱氨酸和蛋氨酸代谢（图4），这两种氨基酸均含有硫。它们可转化为多种有机硫分子，包括N-乙酰半胱氨酸、牛磺酸、维生素B1及辅酶A，从而参与宿主细胞内的硫代谢。同时，通过合成铁-硫（FeS）簇，它们还能维持线粒体代谢功能（Kelly和Pearce，2020）。病毒中存在这些基因，对于维持宿主细胞生长及细胞内酸碱平衡具有关键作用。此外，在营养匮乏的环境中，病毒可能“诱导”宿主激活代谢通路，加速宿主代谢，从而促进病毒复制（Hurwitz和U’Ren，2016）。这不仅促进了海洋微生物环路中碳和能量的流动，也有助于生态平衡的恢复。然而，在已分离的大洋螺菌目相关病毒中鉴定到的AMGs数量和多样性仍相对有限。首个大洋螺菌属（*Oceanospirillum*）噬菌体vB_OliS_GJ44以及首个感染大洋螺菌科（Oceanospirillaceae）的肌尾噬菌体vB_OsaM_PD0307分别仅携带一个AMG，编码MazG-like家族蛋白和TetR家族转录调控蛋白（Zhang等，2021；Zhang等，2023）。

目前关于大洋螺菌目相关病毒的ovHHDDEGs信息仍然有限。在先前的研究中，大洋螺菌目相关病毒的ovHHDDEGs主要来源于Alcanivoracaceae的宿主（Ru等，2023）。本研究中，鉴定出了大量ovHHDDEGs，并发现它们不仅与Alcanivoracaceae相关，还与Endozoicomonadaceae和Halomonadaceae的潜在宿主相关，从而扩展了大洋螺菌目相关病毒的潜在宿主范围。与烃类及烃衍生物降解相关的AMGs中，最频繁HGT关联均出现在大洋螺菌目，包括全部54个AMGs。先前研究中，仅鉴定出15个由病毒携带的烃降解相关的基因（Ru等，2023）。本研究中，鉴定出更多的28个ovHHDDEGs，它们可能参与多种烃降解过程，如二甲苯降解、甲苯降解及其他相关代谢过程（图4和图S3）。此外，其中一些基因还显示出与其他细菌目的HGT关联，提示除了大洋螺菌目外，其他细菌类群也可能获得与烃及烃衍生物降解相关的AMGs（图S4）。推测这些AMGs可能最初由感染大洋螺菌目的病毒通过HGT获得，随后再通过后续的HGT事件转移至新的细菌宿主。病毒与宿主蛋白的结构相似性进一步支持了大洋螺菌目相关病毒与潜在宿主之间频繁HGT的证据，其中与烃降解相关的病毒AMGs可能被转移至潜在宿主，使宿主获得降解烃及烃衍生物的能力。这些与烃降解相关的病毒基因突出了大洋螺菌目相关病毒在生物修复（尤其是石油污染区域）中的潜在作用（Ru等，2023）。从进化与结构的角度来看（图5和图S3），这些蛋白经历了趋同进化，可能是由于面临相似的环境压力或功能需求，例如与相同宿主的相互作用。这种结构相似性很可能与它们在烃及烃衍生物降解中的共享功能相关。先前研究已发现大量参与碳代谢的AMGs，包括通过戊糖磷酸途径（PPP）和三羧酸循环（TCA）进行能量生产（Hurwitz等，2013；Hurwitz和U’Ren，2016）。本研究中对大洋螺菌目相关病毒ovHHDDEGs的鉴定，弥补了之前在碳循环理解中的空白，突出了大洋螺菌目相关病毒在调控微生物群落烃降解能力方面的潜在作用。

4.3　大洋螺菌目相关病毒的生物地理分布

在先前对vB_OsaM_PD0307的研究中发现，大洋螺菌目相关病毒分布于ANT、ARC、EPI和MES，但在BATHY未被检测到（Zhang等，2023），这提示这些病毒感染的宿主可能难以适应深海环境。本研究中，大洋螺菌目相关病毒展现出广泛分布，覆盖所有五个VEZs，这一结果与之前对马里亚纳海沟表层沉积物（深度8900m）中温和噬菌体vB_HmeY_H4907的研究一致（Su等，2023）。本研究中大洋螺菌目相关病毒的丰度在EPI最高，其次为ARC（图6）。在石油泄漏事件中，由于大多数石油产品密度低于海水，油膜通常出现在海面，并在风力和洋流作用下迅速扩散（Tarr等，2016），这导致EPI油污染频率高且范围广。蓝藻自然广泛分布于EPI，同时通过脂肪酸产生烃类，其全球产量的十五烷流量超过海洋总石油输入量的100至500倍（Love等，2021）。这也导致了大洋螺菌目相关病毒在EPI的高丰度。在高丰度病毒中，有三条来源于前病毒，其中ARC比例最高（图6）。由此

推测，温和病毒及整合前病毒可能对生活在这些极端环境中的大洋螺菌目起到关键作用。

5 结论

综上所述，本研究在全球范围内拓展了大洋螺菌目相关病毒的基因组多样性，凸显了大洋螺菌目相关病毒在海洋生态系统中的关键作用，尤其是在HGT、功能多样性及生态适应性方面。与烃及烃衍生物降解相关的大洋螺菌目相关病毒中AMGs的多样性和结构表明，其在烃类降解中的作用可能被忽视。未来，通过从深海、极地海洋等多样且未充分探索的环境中分离更多大洋螺菌目相关病毒，并结合基因组学、宏基因组学及宏转录组学分析，以及对其烃降解功能的验证，将有助于加深我们对大洋螺菌目相关病毒在全球生物地球化学循环中生态潜能的理解。

致谢 感谢崂山实验室(青岛)高性能计算与系统模拟中心提供的高性能服务器支持，以及海洋生物多样性与进化研究所(IEMB-1)高性能计算集群、中国海洋大学高性能生物超算中心以及中国海洋大学海洋高等研究院海洋大数据中心提供的计算资源支持。

补充材料 补充材料为作者提供的原始数据，作者对其学术质量和内容负责，详见网络版(http://earthcn.scichina.com)。

参考文献

Aroney S T N, Newell R J P, Nissen J N, et al. 2025. CoverM: Read alignment statistics for metagenomics. Bioinformatics, 41: btaf147
Bastian M, Heymann S, Jacomy M. 2009. Gephi: An open source software for exploring and manipulating networks. ICWSM, 3: 361–362
Brum J R, Ignacio-Espinoza J C, Roux S, et al. 2015. Patterns and ecological drivers of ocean viral communities. Science, 348: 1261498
Buchfink B, Reuter K, Drost H G. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat Methods, 18: 366–368
Camargo A P, Nayfach S, Chen I M A, et al. 2023. IMG/VR v4: An expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. Nucleic Acids Res, 51: D733–D743
Camargo A P, Roux S, Schulz F, et al. 2024. Identification of mobile genetic elements with geNomad. Nat Biotechnol, 42: 1303–1312
Capella-Gutiérrez S, Silla-Martínez J M, Gabaldón T. 2009. trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics, 25: 1972–1973
Crummett L T, Puxty R J, Weihe C, et al. 2016. The genomic content and context of auxiliary metabolic genes in marine cyanomyoviruses. Virology, 499: 219–229
Darzi Y, Letunic I, Bork P, et al. 2018. iPath3.0: Interactive pathways explorer v3. Nucleic Acids Res, 46: W510–W513
Delmont T O, Eren A M, Vineis J H, et al. 2015. Genome reconstructions indicate the partitioning of ecological functions inside a phytoplankton bloom in the Amundsen Sea, Antarctica. Front Microbiol, 6, doi: 10.3389/fmicb.2015.01090
Edgar R C. 2022. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun, 13: 6968
Gao C, Liang Y, Jiang Y, et al. 2022. Virioplankton assemblages from challenger deep, the deepest place in the oceans. iScience, 25: 104680
Ghiglione J F, Galand P E, Pommier T, et al. 2012. Pole-to-pole biogeography of surface and deep marine bacterial communities. Proc Natl Acad Sci USA, 109: 17633–17638
Gregory A C, Solonenko S A, Ignacio-Espinoza J C, et al. 2016. Genomic differentiation among wild cyanophages despite widespread horizontal gene transfer. BMC Genomics, 17: 1–3
Gregory A C, Zayed A A, Conceição-Neto N, et al. 2019. Marine DNA viral macro- and microdiversity from pole to pole. Cell, 177: 1109–1123.e14
Hurwitz B L, Hallam S J, Sullivan M B. 2013. Metabolic reprogramming by viruses in the sunlit and dark ocean. Genome Biol, 14: 1–4
Hurwitz B L, U'Ren J M. 2016. Viral metabolic reprogramming in marine ecosystems. Curr Opin Microbiol, 31: 161–168
Irwin N A T, Pittis A A, Richards T A, et al. 2022. Systematic evaluation of horizontal gene transfer between eukaryotes and viruses. Nat Microbiol, 7: 327–336
Joglar V, Pontiller B, Martínez-García S, et al. 2021. Microbial plankton community structure and function responses to vitamin B_{12} and B_1 amendments in an upwelling system. Appl Environ Microbiol, 87: e01525–21
Kanehisa M, Furumichi M, Tanabe M, et al. 2017. KEGG: New perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res, 45: D353–D361
Kelly B, Pearce E L. 2020. Amino assets: How amino acids support immunity. Cell Metab, 32: 154–175
Kim H, Choo Y J, Cho J C. 2007. Litoricolaceae fam. nov., to include Litoricola lipolytica gen. nov., sp. nov., a marine bacterium belonging to the order Oceanospirillales. Int J Systatic Evolary Microbiol, 57: 1793–1798
Konstantinidis K T, Tiedje J M. 2005. Genomic insights that advance the species definition for prokaryotes. Proc Natl Acad Sci USA, 102: 2567–2572
Kostka J E, Prakash O, Overholt W A, et al. 2011. Hydrocarbon-degrading bacteria and the bacterial community response in Gulf of Mexico beach sands impacted by the Deepwater Horizon oil spill. Appl Environ Microbiol, 77: 7962–7974
Liang X L, Wang Y S, Zhang Y, et al. 2021. Viral abundance, community structure and correlation with bacterial community in soils of different cover plants. Appl Soil Ecol, 168: 104138
Lin Z M, Akin H, Rao R, et al. 2023. Evolutionary-scale prediction of atomic-level protein structure with a language model. Science, 379: 1123–1130
Liu J L, Xue C X, Wang J Y, et al. 2022. Oceanospirillales containing the DMSP lyase DddD are key utilisers of carbon from DMSP in coastal seawater. Microbiome, 10: 110
Liu J W, Zheng Y F, Lin H Y, et al. 2019. Proliferation of hydrocarbon-degrading microbes at the bottom of the Mariana Trench. Microbiome, 7: 47
Liu J W, Li D W, He X X, et al. 2024. A unique subseafloor microbiosphere in the Mariana Trench driven by episodic sedimentation. Mar Life Sci Technol, 6: 168–181
Liu Z H, Huang Y, Zhang T, et al. 2023. Viruses regulate microbial community assembly together with environmental factors in acid mine drainage. Appl Environ Microbiol, 89: e01973-22
Love C R, Arrington E C, Gosselin K M, et al. 2021. Microbial production and consumption of hydrocarbons in the global ocean. Nat Microbiol, 6: 489–498

Markwitz P, Lood C, Olszak T, et al. 2022. Genome-driven elucidation of phage-host interplay and impact of phage resistance evolution on bacterial fitness. ISME J, 16: 533–542

Middelboe M, Brussaard P C. 2017. Marine viruses: Key players in marine ecosystems. Viruses, 9: 302

Mistry J, Chuguransky S, Williams L, et al. 2021. Pfam: The protein families database in 2021. Nucleic Acids Res, 49: D412–D419

Moore E R, Weaver A J, Davis E W, et al. 2022. Metabolism of key atmospheric volatile organic compounds by the marine heterotrophic bacterium Pelagibacter HTCC1062 (SAR11). Environ Microbiol, 24: 212–222

Nayfach S, Camargo A P, Schulz F, et al. 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nat Biotechnol, 39: 578–585

Nguyen L T, Schmidt H A, von Haeseler A, et al. 2015. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol, 32: 268–274

Roux S, Enault F, Hurwitz B L, et al. 2015. VirSorter: Mining viral signal from microbial genomic data. PeerJ, 3: e985

Roux S, Camargo A P, Coutinho F H, et al. 2023. iPHoP: An integrated machine learning framework to maximize host prediction for metagenome-derived viruses of archaea and bacteria. PLoS Biol, 21: e3002083

Ru J L, Xue J L, Sun J F, et al. 2023. Unveiling the hidden role of aquatic viruses in hydrocarbon pollution bioremediation. J Hazard Mater, 459: 132299

Sabirova J S, Ferrer M, Regenhardt D, et al. 2006. Proteomic insights into metabolic adaptations in *Alcanivorax* borkumensis induced by alkane utilization. J Bacteriol, 188: 3763–3773

Sayers E W, Cavanaugh M, Clark K, et al. 2022. GenBank. Nucleic Acids Res, 50: D161–D164

Schulz F, Roux S, Paez-Espino D, et al. 2020. Giant virus diversity and host interactions through global metagenomics. Nature, 578: 432–436

Shaffer M, Borton M A, McGivern B B, et al. 2020. DRAM for distilling microbial metabolism to automate the curation of microbiome function. Nucleic Acids Res, 48: 8883–8900

Skennerton C T, Imelfort M, Tyson G W. 2013. Crass: Identification and reconstruction of CRISPR from unassembled metagenomic data. Nucleic Acids Res, 41: e105

Steinegger M, Söding J. 2017. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol, 35: 1026–1028

Su Y, Zhang W J, Liang Y T, et al. 2023. Identification and genomic analysis of temperate *Halomonas* bacteriophage vB_HmeY_H4907 from the surface sediment of the Mariana Trench at a depth of 8,900 m. Microbiol Spectr, 11: e01912-23

Sun F L, Wu M L, Wang Y S, et al. 2020. Diversity and potential function of bacterial communities in different upwelling systems. Estuar Coast Shelf Sci, 237: 106698

Tarr M, Zito P, Overton E B, et al. 2016. Weathering of oil spilled in the marine environment. Oceanography, 29: 126–135

van Kempen M, Kim S S, Tumescheit C, et al. 2024. Fast and accurate protein structure search with Foldseek. Nat Biotechnol, 42: 243–246

Jang H B, Bolduc B, Zablocki O, et al. 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nat Biotechnol, 37: 632–639

Zhang W J, Liang Y T, Zheng K Y, et al. 2021. Characterization and genomic analysis of the first *Oceanospirillum* phage, vB_OliS_GJ44, representing a novel siphoviral cluster. BMC Genomics, 22: 675

Zhang W J, Liu Y D, Zheng K Y, et al. 2023. Discovery of an abundant viral genus in polar regions through the isolation and genomic characterization of a new virus against Oceanospirillaceae. Appl Environ Microbiol, 89: e01896-22

Zheng K Y, Dong Y, Liang Y T, et al. 2023a. Genomic diversity and ecological distribution of marine *Pseudoalteromonas* phages. Mar Life Sci Technol, 5: 271–285

Zheng K Y, Liang Y T, Paez-Espino D, et al. 2023b. Identification of hidden N4-like viruses and their interactions with hosts. mSystems, 8: e00197-23

Zheng K Y, Liang, Zhang Y, et al. 2025a. Diversity of reverse-transcriptase-containing viruses through global metagenomics. hLife, 3: 52–56

Zheng K Y, Sun J H, Liang Y T, et al. 2025b. VITAP: A high precision tool for DNA and RNA viral classification based on meta-omic data. Nat Commun, 16: 2226

Zhou K, Zhang R, Sun J, et al. 2019. Potential interactions between clade SUP05 sulfur-oxidizing bacteria and phages in hydrothermal vent sponges. Appl Environ Microbiol, 85: e00992-19

（责任编委：王晓雪）