



专论与综述

## 病毒宏基因组学研究进展

徐志伟 魏云林 季秀玲\*

昆明理工大学生命科学与技术学院 云南 昆明 650500

**摘要:** 病毒宏基因组学是一种新的病毒组学研究手段, 随着高通量测序技术的飞速发展, 人们能够从环境中快速发现、鉴定病毒基因组的组成并研究其特征。在过去的十年里, 研究者们运用病毒宏基因组学发现了许多新型病毒, 增强了人们对不同环境中病毒组成、分布和多样性的了解。因此, 病毒宏基因组学已成为清晰描绘各种特殊环境中病毒图谱、了解自然界中病毒分布动态的有效工具。本文主要从病毒宏基因组的概念、样品前处理和病毒总基因组提取方法、测序技术以及病毒宏基因组的应用和发展前景方面进行概述。

**关键词:** 病毒宏基因组, 环境微生物, 高通量测序, 新病毒鉴定

## Advances in viral metagenomics

XU Zhi-Wei WEI Yun-Lin JI Xiu-Ling\*

Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming, Yunnan 650500, China

**Abstract:** Viral metagenomics is a new approach to genomics research. With the rapid development of high-throughput sequencing technology, viral genomes from the environment can be quickly discovered, identified and described based on composition characteristics. In the past decade, many novel viruses have been discovered by viral metagenomics, enhancing the understanding of virus composition, distribution and diversity in different environments. Therefore, viral metagenomics has become an effective tool for clearly depicting the viral map in various special environments and understanding the distribution dynamics of viruses in nature. This article mainly summarizes the concept of virus metagenome, sample preparation and extraction method of virus total genome, sequencing technology, the application and development prospects of virus metagenome.

**Keywords:** Viral metagenomics, Environmental microbes, High throughput sequencing, New virus identification

病毒是地球上最为丰富的微生物资源。病毒个体微小, 多数直径在 18–450 nm 之间, 结构简单, 只含有一种核酸(DNA 或 RNA), 是必须在活

细胞内寄生并以复制方式增殖的非细胞型生物。病毒是地球上最丰富的生物实体, 影响着微生物群落的结构和功能<sup>[1]</sup>。目前开发的病毒数量尚未达

**Foundation item:** National Natural Science Foundation of China (31700324)

**\*Corresponding author:** Tel: 86-871-65920147; E-mail: jixiuling1023@126.com

**Received:** 27-09-2019; **Accepted:** 26-04-2020; **Published online:** 28-05-2020

基金项目: 国家自然科学基金(31700324)

**\*通信作者:** Tel: 0871-65920147; E-mail: jixiuling1023@126.com

收稿日期: 2019-09-27; 接受日期: 2020-04-26; 网络首发日期: 2020-05-28

到 1%; 加上对其基因组成的认识有限, 尽管病毒对整个生物学研究而言具有非常重要的意义, 但由于捕获完整病毒基因组的方法受到科学技术的限制以及难以产生足够量的基因组用于测序, 因此对病毒的研究始终无法深入。随着科技的不断进步, 宏基因组学的提出为解决这一难题提供了良好的方案。病毒宏基因组学是指从环境或者生物组织中富集和提取病毒基因组后进行生物信息学分析, 由于高通量测序技术的发展克服了环境样本中病毒浓度低和基因组易受干扰的问题, 从而使得病毒宏基因组学能够更广泛地运用到各种复杂的环境中<sup>[2]</sup>。病毒宏基因组学揭示了未开发病毒的遗传多样性, 同时加深了人们对病毒与生态环境和宿主之间相互作用的认识和理解。

## 1 病毒宏基因组学

宏基因组的概念首次由 Handelsman 等 1998 年提出, 其定义为“the genomes of the total microbiota found in nature”, 即环境中全部微小生物遗传物质的总和, 包括可培养和不能培养的微生物; 其优势是能够绕过微生物的分离和培养过程, 成为研究环境样品中超过 90% 不可培养微生物的有力手段<sup>[3]</sup>。宏基因组学(metagenomics)以环境样品中的微生物群体基因组为研究对象, 结合了分子生物学的研究方法, 以功能基因筛选和测序分析为研究手段, 以微生物多样性、种群结构、进化关系、功能活性、相互协作关系及与环境之间的关系为研究目的, 成为一种新的微生物研究方法<sup>[4]</sup>。一般流程包括从环境样品中提取总基因组 DNA、克隆 DNA 到合适的载体、将载体转化到宿主细胞内并构建基因组文库、筛选目的转化子等工作<sup>[5]</sup>。

如果将宏基因组学的研究方法应用到病毒领域, 就形成了病毒宏基因组学(viral metagenomics)<sup>[6]</sup>。2002 年 Breitbart 等开展了海水中病毒组的研究, 为海洋病毒宏基因组学的正式开展拉开了帷幕<sup>[7]</sup>。病毒宏基因组的研究手段是将环境中的病毒与其他微生物分离, 只研究病毒的

基因组, 然后将得到的病毒基因组数据同目前已知的病毒数据进行比对和分析, 可以直接研究不同环境中病毒的一些变异情况, 也能防控追踪和预测新型的病毒; 特别是对于那些特殊结构的病毒, 病毒宏基因组分析将提供关于病毒遗传和表型多样性的关键数据, 改变人们对以往认识的病毒圈的理解<sup>[8]</sup>。

## 2 病毒宏基因组学的研究过程

病毒宏基因组学的研究过程主要包括 3 个步骤: 样品的前处理; 病毒宏基因组学文库构建; 高通量测序与数据分析<sup>[9]</sup>。图 1 列出了病毒宏基因组和其他微生物的宏基因组样品的处理流程, 可以看出这两种方法的基本处理流程是一样的, 最大的区别就在于前期样品的预处理方式不同。由于病毒基因组序列相对较短, 细菌或真核生物的核酸会严重干扰病毒基因组的分离和检测。因此, 在前期样品预处理的过程中需要去除非病毒核酸, 通常采用的方式是先用氯仿处理再加 DNase 消化以去除污染的 DNA。氯仿能破坏细菌和真核生物的细胞膜, 暴露非病毒 DNA, 然后被 DNase 消化。去除非病毒核酸后再用均质化、过滤和超速离心的方法浓缩样品中的病毒颗粒<sup>[10]</sup>。

### 2.1 样品前处理及病毒总基因组提取方法

环境中样品 DNA 的质量直接关系到病毒宏基因组分析的结果<sup>[11]</sup>。在开展病毒宏基因组分析时样品的选择和处理非常重要, 应该选择病毒含量相对丰富且容易处理的样品。对于含水样品, 可以选用切向流过滤(tangential flow filtration, TFF), 例如古氏病毒科(一种古细菌病毒家族)的病毒粒子能在氯化铯中溶解, 可以采用 TFF 过滤后用蔗糖梯度进行纯化<sup>[12]</sup>。TFF 浓缩病毒样品的缺点是容易受到细菌基因组的干扰<sup>[13]</sup>。对于非水类样品, Thurber 等<sup>[12]</sup>比较了 4 种来自粪便样品的纯化病毒样颗粒(virus-like particles, VLP)的方法, 结果显示: CsCl 密度梯度离心在去除宿主衍生的 DNA 方面有效, 但对特异性噬菌体具有强烈区

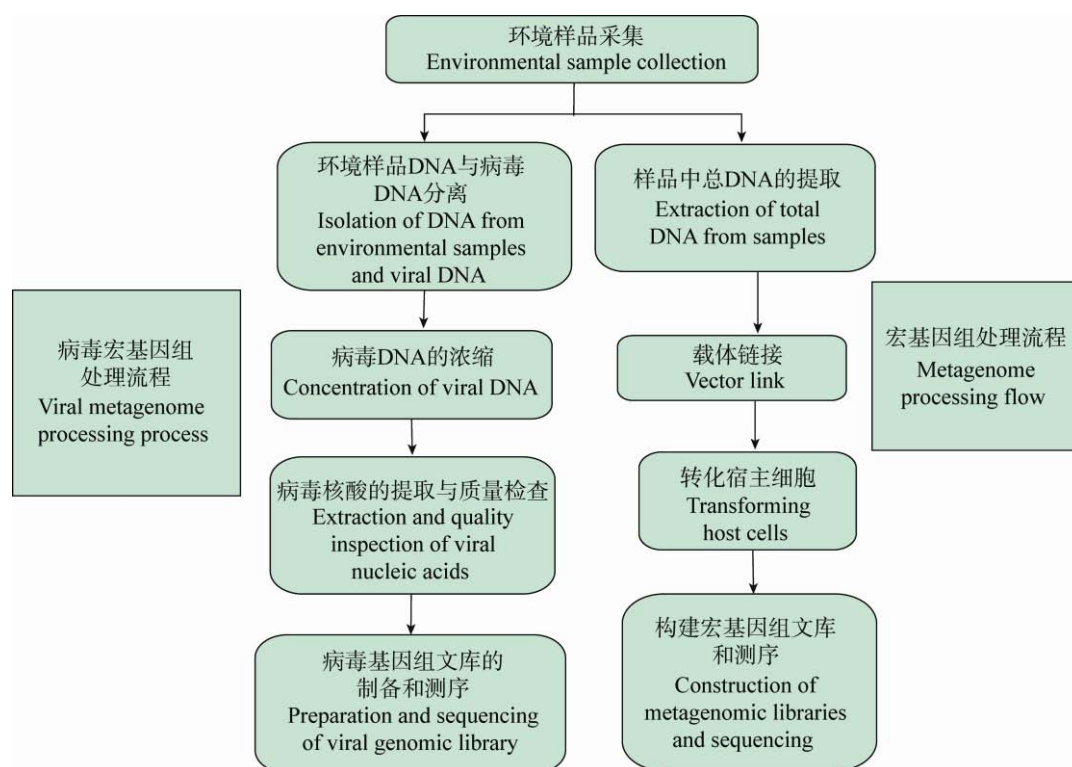


图1 病毒宏基因组和其他微生物宏基因组样品处理流程

Figure 1 Viral metagenomic and other microbial metagenomic sample processing procedures

分并且不利于定量分析。浓缩过后的病毒样品可以进行核酸提取，提取核酸主要有两种方式：第一种是试剂盒法，比如 QIAamp MinELute Virus Spin Kit，这类试剂盒可用于 DNA 病毒的提取并且具有方便、提取纯度高等优点；第二种是利用核酸、蛋白等杂质在水相和有机相中溶解性的不同而重新分配，常用的是酚-氯仿提取法，该方法经济廉价，但提取效率不高且纯度往往不够<sup>[14]</sup>。提取的 DNA 如果浓度达不到测序要求，可以进行核酸扩增，利用 Genomiphi DNA 试剂盒扩增 DNA，得到的 DNA 可用于构建基因文库或高通量测序。如果样品有 RNA 的污染，可以通过加入 RNase 以消除 RNA 的干扰<sup>[15]</sup>。但如果环境样品是 RNA 病毒，那么获得的 RNA 基因组只需要去除宿主 DNA 和环境中的 DNA 病毒。通常的做法是先利用 0.22  $\mu\text{m}$  或者 0.45  $\mu\text{m}$  的滤膜选择性富集环境中的基因组，然后加入 DNase I 或者 Benzonase 处理

30 min 即可去除所有干扰的 DNA 而留下 RNA 病毒。留下来的高浓度 RNA 病毒利用病毒宏基因组学可以恢复与复制 DNA 病毒相关的 mRNA。

## 2.2 构建基因组文库

构建病毒宏基因组文库时要考虑到自身的研究目的和获得的 DNA 的量、纯度及片段大小等因素。另外，DNA 片段的大小决定了基因组文库完整性<sup>[16]</sup>。目前构建宏基因组文库有两种方式，分别是载体克隆和基于新一代测序技术的加接头文库。常用于构建文库的载体有质粒(plasmid)、柯斯载体(cosmid)、福斯质粒(fosmid)和细菌人工染色体(bacterial artificial chromosome, BAC)等，能够满足不同插入片段大小的要求。载体的选择也要根据插入片段的大小进行选择，例如：DNA 片段 <10 kb，适合插入质粒载体；DNA 片段 30–45 kb，适合黏粒(fosmid/cosmid)；DNA 片段 >40 kb，适合细菌人工染色体(bacterial artificial

chromosome, BAC)<sup>[17]</sup>。新一代测序技术的加接头文库以 SOLiD 为例, SOLiD 全称为 Supported oligo ligation detection, 它是以四色荧光标记寡核苷酸的连续连接合成为基础, 取代了传统的聚合酶连接反应, 可对单拷贝 DNA 片段进行大规模扩增和高通量测序。SOLiD 支持片段文库和配对末端文库。片段文库是将基因组 DNA 打断, 两头加上接头, 制成文库, 而配对末端文库则是将基因组 DNA 打断后与中直接头连接, 再环化, 然后用 *EcoP15* 酶切, 使中直接头两端各有 27 bp 的碱基, 再加上两端的接头形成文库<sup>[18]</sup>。在实际应用中可以根据研究目的以及基因组学文库的筛选方法来选择合适的方法构建基因组文库。

### 2.3 基因组文库的筛选

基因组文库的筛选通常有 3 种方法: 序列筛选(sequence-based screening)、功能筛选(functional screening)和底物诱导基因筛选(substrate-induced gene expression screening, SIGEX)<sup>[19]</sup>。序列筛选是根据已知的功能基因或蛋白保守区域设计探针或者 PCR 引物, 通过杂交或 PCR 扩增筛选阳性克隆子; 优点是不依赖宿主菌来表达目的基因, 将 DNA 分子杂交技术和 PCR 扩增技术应用于筛选工作, 提高了筛选效率, 能够筛选到某一类新结构或功能的蛋白质分子<sup>[20]</sup>。功能筛选是针对具有特殊功能的克隆子进行检测, 可以从多个克隆子中鉴定出全长基因并获得其功能产物, 灵敏度高。现如今医药和农业的一些天然活性物质, 如蛋白酶、纤维素酶和脂肪酶等都是通过这种方法筛选出功能基因的产物。底物诱导基因筛选是利用合适的底物诱导克隆子分解代谢基因表达来进行筛选。Daniel 将外源基因插入到 *lac* 启动子和 *gfp* 绿色荧光蛋白基因之间, 利用底物诱导基因筛选的方法筛选出 35 个目标克隆, 其优点是不需要对底物进行修饰, 可以从底物推断出相应的酶, 从而推断基因的功能, 为高通量筛选提供了保障<sup>[21]</sup>。

### 2.4 高通量测序技术

宏基因组学获得的各种基因资源依赖于高通量测序技术的发展。最早的大规模测序工作是 Venter 等对 Sargasso Sea 环境基因组的研究, 应用鸟枪法对 Sargasso Sea 表层水样的基因组文库进行测序, 结果发现大量新的基因和物种, 大大扩展了对海洋微生物多样性的认识<sup>[22]</sup>。随后由于测序技术的不断进步, 出现了二代测序技术(454 焦磷酸测序)以及三代测序技术(单分子实时 DNA 测序)。目前应用于病毒宏基因组的测序技术主要是二代和三代测序技术。随着测序技术的高速发展, 病毒宏基因组学数据越来越丰富。高通量测序的结果需要使用专门的软件来处理, 例如 RAMMCP 和 MetaStats, 然后将所得测序结果和已知的病毒数据库进行比对, 确定该序列的生物群落来源, 筛选出有用的基因信息。此外, 还可以用一些工具对序列进行基因预测(如 Metagenemark 和 BLASTp 等)。

## 3 病毒宏基因组学的应用

### 3.1 土壤病毒宏基因组学

病毒宏基因组学以其显著的优势在病毒研究的各个方面得到广泛应用, 例如人类和动植物健康、水环境问题、病毒领域的研究等, 改变着人们对病毒的认识, 也为解决病毒相关问题提供了新的思路。土壤是微生物的大本营, 同时含有丰富的病毒资源。病毒宏基因组技术在土壤方面的应用主要包括两方面: 一是土壤微生物及其病毒资源的挖掘, 另一方面是揭示土壤微生物和病毒的多样性及其与环境之间的关系。

Paez-Espino 等利用 3 042 个地理上不同样本的病毒基因组数据评估了病毒的全球分布、系统发育多样性和宿主特异性<sup>[23]</sup>, 研究发现超过 125 000 个 DNA 病毒基因组, 把已知病毒基因的数量提高了 16 倍。Adriaenssens 等第一次提出利用病毒宏基因组学的方法研究纳米布沙漠中的病毒群落, 了解了纳米布沙漠整体病毒群落的分类组

成, 并完成了序列的功能评估<sup>[24]</sup> (表 1)。他们还分析了南极大陆 14 个土壤样本病毒群落的病毒宏基因组数据, 发现病毒群落组成的环境驱动因素和高度无关, 只与土壤的 pH 及钙含量有关<sup>[25]</sup>。季秀玲<sup>[26]</sup>针对纳帕海高原湿地开展了噬菌体丰度、时空分布、遗传多样性研究, 并探讨其在该区域碳循环中的作用和功能; 在时空分布上旱季和雨季浮游细菌丰度的平均值分别为  $3.52 \times 10^5$  cells/mL 和  $2.02 \times 10^6$  cells/mL, 浮游病毒丰度的平均值分别为  $3.63 \times 10^6$  cells/mL 和  $3.71 \times 10^7$  cells/mL, 而且雨季浮游病毒和细菌的丰度均显著高于旱季, 另外对有机碳释放量和纳帕海湿地中可溶性有机碳的分布和数量进行了测定, 通过经典的数学模型对二者的关

系进行了研究。研究结果表明在雨季细菌的有机碳产出量为  $8.01 \mu\text{g-C}/(\text{L} \cdot \text{h})$ , 旱季为  $10.3 \mu\text{g-C}/(\text{L} \cdot \text{h})$ , 在雨季和旱季经由病毒裂解而引起的细菌死亡率 (fraction of mortality from viral lysis, FMVL) 分别为 54.9% 和 27%; 相关数据通过 Binder 数学模型建模分析后表明在纳帕海地区经由病毒裂解细菌所贡献的可溶性有机碳在总有机碳中占主导地位, 其贡献率无论在旱季和雨季均超过 50%, 并且在雨季和旱季其贡献率有显著差异, 分别为 84.37% 和 72.9%。张中耀、王爽以纳帕海高原湿地分离得到的低温噬菌体为材料, 开展了噬菌体-宿主共进化方面的研究, 对不同进化时期的宿主和噬菌体进行了全基因组解析, 并确定了噬菌体的吸附元件<sup>[27-28]</sup>。

表 1 来自纳米布沙漠病毒组 SEED 类别和相关基因的功能预测<sup>[24]</sup>

Table 1 Functional prediction of SEED categories and related genes from the Namib Desert virus group<sup>[24]</sup>

SEED 功能类别	纳米布沙漠病毒组中预测基因	预测功能
SEED functional category	Predicted genes in Namib hypolith metavirome	Putative function
鼠李素水解酶	N-乙酰基-壁酰-L-丙氨酸酰胺酶, 尾纤维羧肽酶	细菌细胞壁降解
Murein hydrolases	N-acetyl-muramoyl-L-alanine-amidase, tail fiber, Carboxypeptidase	Bacterial cell wall degradation
肽聚糖氨基酸的循环利用	N-乙酰基-壁酰-L-丙氨酸酰胺酶, 尾纤维	细菌细胞壁降解
Recycling of peptidoglycan amino acids	N-acetyl-muramoyl-L-alanine-amidase, tail fiber	Bacterial cell wall degradation
基于聚类的子系统	噬菌体乙酰壁酰丙氨酸酰胺酶(金属内肽酶), 产孢蛋白	细菌细胞壁降解
Cbss-393121.3.peg.2760	Phage N-acetyl-muramoyl-L-alanine amidase, (metalloendo) peptidase, sporulation protein	Bacterial cell wall degradation
Clustering-Based Subsystems		
Cbss-393121.3.peg.2760		
锌依赖的细菌 RNA 代谢水解酶类	细胞分裂蛋白	DNA 转移
Bacterial RNA-metabolizing	Cell division protein, FtsK/SpoEIII	DNA transfer
Zn-dependent hydrolases		
噬菌体尾部蛋白 2	尾带测量蛋白质, 主要尾部蛋白	噬菌体颗粒组装
Phage tail proteins 2	Tail tape measure protein, major tail protein	Phage particle assembly
R1t 样链球菌噬菌体	结构蛋白, 核酸内切酶, 噬菌体相关蛋白, 核酸内切酶, 内肽酶,	噬菌体颗粒组装
R1t-like streptococcal phages	假定蛋白	Phage particle assembly
	Structural proteins, homing endonucleases, phage-associated proteins, endonuclease, endopeptidase, hypothetical protein	
李斯特 phi-a118 样前噬菌体	假定的尾蛋白或底板蛋白, 次要/主要衣壳蛋白, terS, 复制蛋白,	噬菌体颗粒组装
Listeria phi-a118-like prophages	赖氨酸, tmp, 假定蛋白, ssDNA 结合蛋白, 阻遏物, 抗阻遏物	Phage particle assembly
	Putative tail or baseplate protein, minor/major capsid protein, terS, replication protein, lysin, tmp, hypothetical protein, ssDNA binding protein, repressor, antirepressor	
磷酸盐代谢	假定蛋白	未知功能
Phosphate metabolism	Hypothetical protein, phoH	Unknown function
核糖核苷酸还原	核糖核苷酸还原酶	dNTP 合成
Ribonucleotide reduction	Ribonucleotide reductases	dNTP synthesis

为了更加深入地了解纳帕海高原湿地病毒的多样性, 季秀玲等从纳帕海高原湿地 6 个土壤样本的病毒宏基因组数据得到了病毒科水平的分类和注释(图 2), 发现剑叶病毒科(Siphoviridae)和基因病毒科(Genomoviridae)这两类病毒占病毒样本的 60%以上, 还有一些尚未分类的病毒占 20%的比例, 表明纳帕海高原湿地土壤中存在一些新的病毒种群。借助病毒宏基因组方法不仅可以系统研究该环境中病毒样品的群落分布, 还可以在基因层面对病毒的基因组进行注释和分析, 更加全面地了解病毒的遗传多样性。病毒宏基因组学成功运用到土壤环境的研究中, 扩大了现有病毒数据库, 了解了病毒群落的分类、组成以及环境驱动因素, 拓展了对土壤生态环境中病毒群落的认识。

### 3.2 海洋病毒宏基因组学

海洋具有极为丰富的病毒资源, 病毒宏基因组的出现为海洋病毒基因组的研究带来了便捷, 将人们对病毒的视野延伸到了以往无法探测的未知领域, 对病毒的海洋生态作用有了更深入的了解。海洋表层中的病毒表现出丰富的多样性, 海水中病毒颗粒含量约为  $10^7$  cells/mL, 这些病毒在其宿主的发育、进化中起到不可或缺的作用。Nishimura 等<sup>[29]</sup>应用病毒宏基因组学的方法分析了由 52 个海洋病毒体组装而成的 1 600 个双链 DNA 病毒基因组, 确定了 1 352 个非冗余完整病毒基因

组。通过对这些病毒基因组的聚类分析和预测, 揭示出海洋病毒-宿主相互作用的新机制: 这些海洋噬菌体通过编码一些特殊的酶来增强对铁硫簇的有效利用或增强宿主的存活率, 该研究不仅发现了数百个新型病毒属, 还表明海洋病毒对一些基因是有选择的, 从而积累有利于病毒繁殖的基因来控制宿主细胞的代谢。海洋中的病毒通过影响微生物种群的大小、多样性等来影响生态系统, 那么我们对于海洋病毒群落的分布模式和局部多样性了解多少呢? 借助病毒宏基因组学, Brum 等<sup>[30]</sup>分析了来自 43 个 Tara Oceans 海洋病毒样本的宏基因组数据研究海洋病毒种群模式和结构, 并从病毒蛋白簇、种群和病毒形态的分析解释了海洋病毒种群较高的局部多样性和较低的全局多样性, 该研究提供了全球上层海洋病毒群落大规模的数据以及图片, 为今后病毒生态学模拟驱动地球生态系统(病毒级)提供了基础。海洋病毒无处不在, 但是我们对于病毒在海底生物圈的作用知之甚少。运用病毒宏基因组学手段, Cai 等对波罗的海沉积物中的病毒进行了定量和表征, 结果显示波罗的海海底生物圈中病毒丰度高达  $1.8 \times 10^{10}$  viruses/cm<sup>3</sup>, 并且首次在深海海底的原核细胞中观察到完整的病毒样颗粒, 证明了宿主体内病毒颗粒的原位组装过程, 表明了病毒通过影响海底微生物群落的动态, 从而影响海底生物圈中的生物地球化学过程<sup>[31]</sup>。

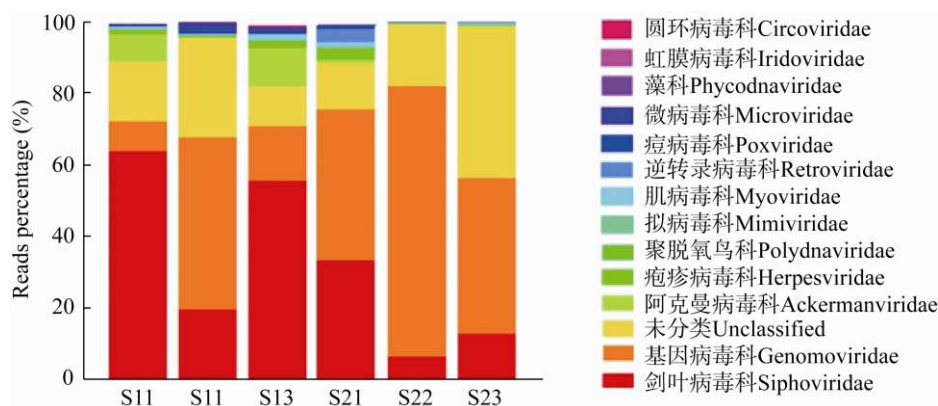


图 2 纳帕海病毒在科水平注释结果统计图

Figure 2 Statistical chart of annotation results of Napahai virus at family level

3.3 动物疾病和肠道健康与病毒宏基因组学

将病毒宏基因组学运用到有机环境中，利用病毒宏基因组学的手段可以对病毒快速鉴定，并了解这些威胁人和动物健康的病毒组的构成，从而实现对一些传染病及时的防控。Ng 等<sup>[32]</sup>富集并测序了 50 只具有牛呼吸道疾病(bovine respiratory disease, BRD)症状的年轻奶牛病毒样本，将病毒宏基因组方法与相关案例对照结合后，利用

Illumina HiSeq 平台进行病毒宏基因组学处理，发现这些病毒的数量依次为：牛鼻炎病毒>牛腺病毒>腺相关病毒>牛鼻炎 B 病毒>星状病毒 BSRI1>D 型牛流感病毒>微小疱疹病毒(PBV)>牛细小病毒>牛疱疹病毒，这些病毒的 GenBank 登录号见表 2；通过与病例呼吸道中的病毒匹配，研究者确认了 3 种与牛呼吸道疾病相关的病毒，分别是 Bovine adenovirus 3、Bovine rhinitis A virus 和 Bovine

表 2 病毒宏基因组学鉴定的病毒<sup>[32]</sup>

Table 2 Viruses identified by virus metagenomics<sup>[32]</sup>

病毒 Virus	科 Family	GenBank 登录号 GenBank accession number	片段长度 Contig length (bp)	序列读取数 Number of sequence reads	近亲 Closest relative	氨基酸同一性 Amino acid identity (%)
牛腺病毒 3 Bovine adenovirus 3	腺病毒科 Adenoviridae	KP264982	2 547	4 279	牛腺病毒 3 Bovine adenovirus 3	98 (Hexon)
牛腺相关病毒 Bovine adeno-associated virus	细小病毒科 Parvoviridae	KP264981	4 279	3 481	牛腺相关病毒 Bovine adeno-associated virus	98 (Rep)
牛鼻炎病毒 Bovine rhinitis A virus	鸟苷科 Picornaviridae	KP264974	7 207	11 109	牛鼻炎 A 病毒 Bovine rhinitis A virus	94 (polyprotein)
牛鼻炎 B 病毒 Bovine rhinitis B virus	鸟苷科 Picornaviridae	KP264980	7 359	1 224	牛鼻炎 B 病毒 Bovine rhinitis B virus	96 (polyprotein)
BSRI1	鸟苷科 Picornaviridae	KP264976 to KP264979	1 527	985	牛鼻炎 B 病毒 Bovine rhinitis B virus	92 (3D)
BSRI2	鸟苷科 Picornaviridae	KP264975	6 995	430	牛鼻炎 B 病毒 Bovine rhinitis B virus	89 (polyprotein)
牛流感病毒 Bovine influenza D virus	正粘病毒科 Orthomyxoviridae	KP264970	6 099	304	猪流感 D/OK 病毒 Swine influenza D/OK virus	92–98 (multiple)
牛星形病毒 BSRI1 Bovine astrovirus BSRI1	星形病毒科 Astroviridae	KP264972	1 286	21	牛星形病毒 K08/51 Bovine astrovirus K08/51	71 (RdRp)
小核糖核酸病毒 Picobirnavirus	小核糖核酸病毒 Picobirnaviridae	KP264973	1 366	123	人小核糖核酸病毒 Human picobirnavirus	33 (RdRp)
BSRI1	小核糖核酸病毒 Picobirnaviridae	KP264971	5 331	79	人小核糖核酸病毒 Human picobirnavirus	34 (RdRp)
牛细小病毒 2 型 Bovine parvovirus 2	小核糖核酸病毒 Picobirnaviridae				牛细小病毒 2 Bovine parvovirus 2	96 (NS1)
BSRI	疱疹病毒科 Herpesviridae			19	牛疱疹病毒 6 型 Bovine herpesvirus 6	96 to 100 (multiple)
牛疱疹病毒 6 型 Bovine herpesvirus 6						



influenza D virus, 并且这 3 种病毒在 BRD 病例中的检出率为 62%。呼吸道疾病是对养殖业经济影响最大的疾病, 但影响因素包括环境、遗传和传染等, 确定致病病毒比较困难。Klima 等<sup>[33]</sup>通过病毒宏基因组学方法对 BRD 病例的呼吸道分泌物中的病毒直接进行了表征, 再结合病例对照可以提供与复杂传染病相关的候选病原体, 将研究的范围有效地缩小, 为进一步的研究提供参考。例如 Sachsenröder 等利用病毒宏基因组学在城市野生老鼠的肠道分析鉴定出一种 Group A 轮状病毒并对其特征进行了分析和描述, 该病毒的序列对人和动物显示出高度的同源性, 这对其将来向人类传播的能力提供了有效评估<sup>[34]</sup>。Fawaz 等<sup>[35]</sup>从 23 只鸭子的泄殖腔中提取病毒基因组, 通过病毒宏基因组学的方法获得了鸭肠道病毒数据库, 并对这些真核病毒进行了分类(图 3), 发现 7.33%的序列与

噬菌体相似, 而 2.21%的序列与真核病毒家族成员相似, 包括拟真病毒科、分体病毒科、痘病毒科、藻类去氧核糖核酸病毒科、细小病毒科、通孔病毒科、逆转录病毒科和双顺反病毒科等, 涵盖了无脊椎动物、脊椎动物和植物病毒, 他们通过针对病毒参考数据库的重叠群和单体的序列相似性鉴定了感染鸭的病毒体。

病毒宏基因组的出现为一些高传染性和高致死性疾病提供了研究的可行性, 利用与病毒宏基因组结合开发的新技术还可以对病毒进行实时监测, 为人类未来战胜这些疾病奠定基础。此外, 肠道中病毒是最丰富、最为多样的生物, 其对肠道微生物菌群的调节起到关键作用。病毒宏基因组学方法也可以对这些肠道中的病毒进行研究。Young 等通过病毒宏基因组的分析, 揭示了肺移植接受者体内呼吸道内环状腺病毒的特征, 并发现

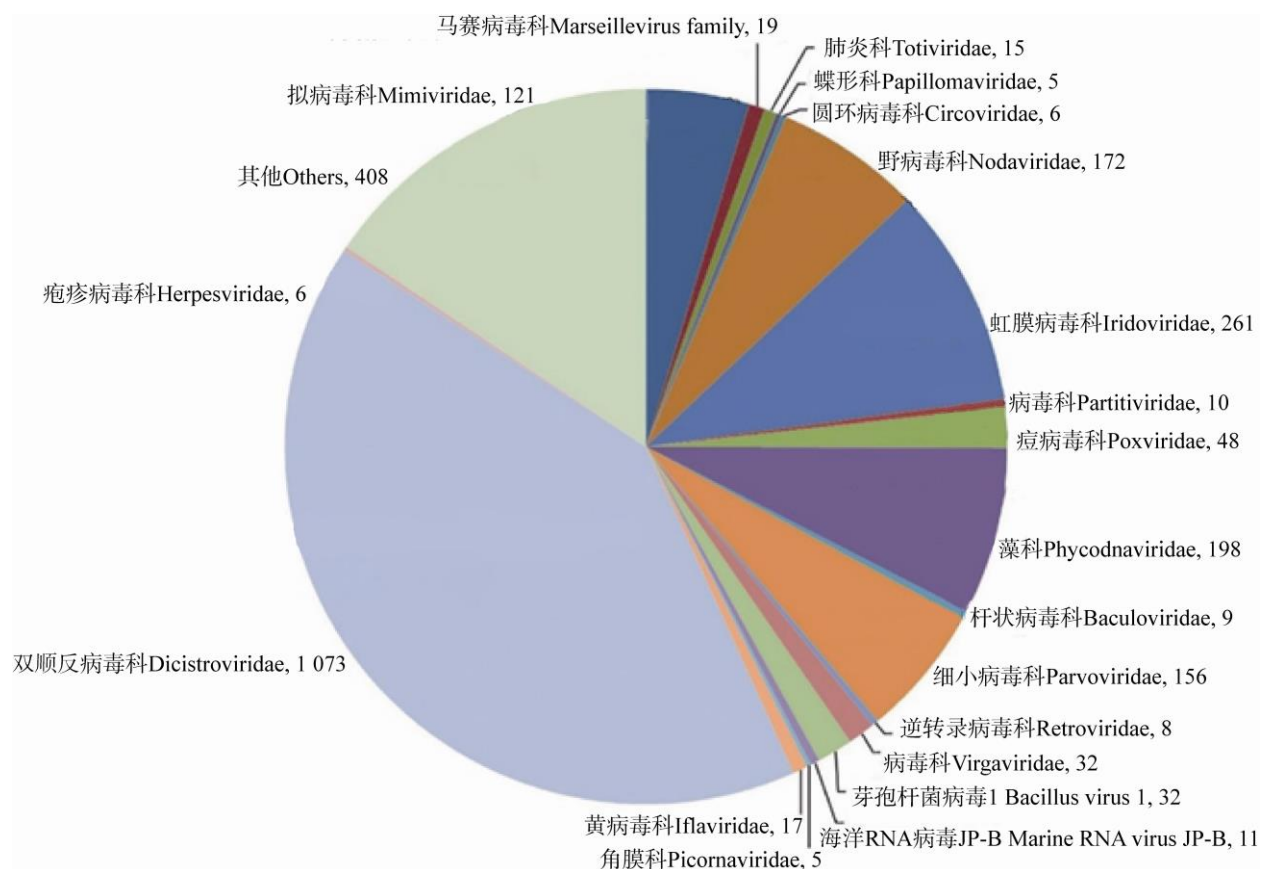


图 3 鸭肠道真核病毒相关序列的分类学分布<sup>[35]</sup>

Figure 3 Taxonomic distribution of eukaryotic virus-related sequences in duck intestines<sup>[35]</sup>



同种异体病毒的载量与同种异体移植支气管肺泡(bronchoalveolar allograft lavage, BAL)中的细菌群落相关<sup>[36]</sup>。Arkhipova 等分析人体肠道内病毒宏基因组,发现了人体肠道的噬菌体微生物组对人体肠道稳态的重要性,病毒体是人体肠道微生物群变化最大的成分之一<sup>[37-38]</sup>。Moreno-Gallego 等发现在成人单卵双胞胎的体内,随着双胞胎年龄的增长,肠道环境中的病毒体基因组也会随之变化产生差异,这些差异是病毒在快速变化的生态系统中能够适应的一个重要因素<sup>[39]</sup>。噬菌体基因组的深度测序揭示了病毒群体在调节肠道整体微生物组成的潜在作用。实际上,肠道噬菌体早就成为一种有效的治疗策略和诊断方法,近年来病毒宏基因组在人体临床检查中也得到了广泛应用。

#### 4 前景与展望

传统方法只能以已知病毒为目标,很难发现新的病毒,而病毒宏基因组学的方法则结合新一代的测序技术和随机 PCR 技术,能够挖掘海量的未知病毒,且不需要分离培养,对一些过于分散、丰度低的病毒也可以进行系统的分析和鉴定。中国拥有非常丰富的生物资源,新病毒的发掘潜力巨大,病毒宏基因组学在这一方面拥有独特的优势。科学家们借助病毒宏基因组学方法已经成功地从环境样品中鉴定出多种病毒,使人们了解到不同环境的病毒构成和一些中间宿主的病毒携带情况。这不仅丰富了病毒库,还增加了对病毒分布和多样性的了解,从而更能科学地防治病毒的侵染。此外,利用病毒宏基因组学实时监控一些致病性病毒和潜在致病性病毒的变化情况,也有助于诊断一些新发或尚不明病原体的疾病,为快速侦检突发病毒性公共卫生事件提供有效的技术手段和方法。

病毒宏基因组学具有广泛的应用前景,但是在整个技术流程中也存在一些技术问题。如:  
(1) 样本基因组的提取方式问题。在利用这项技术之前需要将病毒同其他杂质进行分离过滤,每一次的处理都不可避免地损失很多的病毒粒子。

有时加入核酸酶来排除游离核酸的影响,但这种核酸酶消化的效果无法保证 100%消化游离核酸,因此在病毒样品中存在很多无关序列,这些无关序列不仅增加了测序的成本,还为后续分析增加了难度。(2) 基因文库的筛选方法也需要进一步的完善和改进。

未来在研究极端环境中病毒的结构和应用方面,病毒宏基因组学将成为有力工具。目前,极端环境中病毒功能探索仅限于基于序列的筛选。随着宏蛋白组学、宏转录组学等新技术的加入,病毒宏基因组学的方法能更广泛和准确地应用于病毒相关的领域。总之,病毒宏基因组学技术必将突破传统技术上的瓶颈,在未来展现其独特的优势。

#### REFERENCES

- [1] Zhao R, Hu YF, Jin Q. Macro-genomics and its application in medical microbiology[J]. Chinese Journal of Virology, 2009, 25(3): 231-234 (in Chinese)  
赵蓉, 胡永峰, 金奇. 宏基因组学及其在医学微生物学领域的应用[J]. 病毒学报, 2009, 25(3): 231-234
- [2] Fan ST, Gao YW, Xia XZ. Application of viral metagenomics in medical field[J]. Chinese Journal of Biologicals, 2014, 27(2): 285-288 (in Chinese)  
范胜涛, 高玉伟, 夏咸柱. 病毒宏基因组学在医学领域的应用[J]. 中国生物制品学杂志, 2014, 27(2): 285-288
- [3] Yin L, Gao XD, Gu JF. Research progress of macro-genomic technology[J]. Chinese Medicinal Biotechnology, 2012, 7(3): 216-220 (in Chinese)  
印蕾, 高向东, 顾觉奋. 宏基因组技术研究进展[J]. 中国医药生物技术, 2012, 7(3): 216-220
- [4] Wang K, Wang SD, Huang R, et al. Advances of metagenomics in discovering novel biocatalysts[J]. Chinese Journal of Biotechnology, 2012, 28(4): 420-431 (in Chinese)  
王魁, 汪思迪, 黄睿, 等. 宏基因组学挖掘新型生物催化剂的研究进展[J]. 生物工程学报, 2012, 28(4): 420-431
- [5] Song PY, Ma LL, Wang QR, et al. The advances of metagenome technology and its applications[J]. Guizhou Agricultural Sciences, 2009, 37(10): 14-18 (in Chinese)  
宋培勇, 马莉莉, 王庆容, 等. 宏基因组技术及其应用研究进展[J]. 贵州农业科学, 2009, 37(10): 14-18
- [6] He B, Tu CC. The advances and applications of viral metagenomics[J]. Acta Veterinaria et Zootechnica Sinica, 2012, 43(12): 1865-1870 (in Chinese)  
何彪, 涂长春. 病毒宏基因组学的研究现状及应用[J]. 畜牧兽医学报, 2012, 43(12): 1865-1870

- [7] Breitbart M, Salamon P, Andresen B, et al. Genomic analysis of uncultured marine viral communities[J]. Proceedings of the National Academy of Sciences of the United States of America, 2002, 99(22): 14250-14255
- [8] Zhang YZ, Shi M, Holmes EC. Using metagenomics to characterize an expanding virosphere[J]. Cell, 2018, 172(6): 1168-1172
- [9] Gu W, Miller S, Chiu CY. Clinical metagenomic next-generation sequencing for pathogen detection[J]. Annual Review of Pathology: Mechanisms of Disease, 2019, 14(1): 319-338
- [10] Zhang D, Lou XY, Yan H, et al. Metagenomic analysis of viral nucleic acid extraction methods in respiratory clinical samples[J]. BMC Genomics, 2018, 19(1): 773
- [11] Chen JH. Progress in the study of macrogenomics of environmental microorganisms[J]. Jilin Agricultural, 2010(6): 44-45, 49 (in Chinese)  
陈金华. 环境微生物的宏基因组学研究进展[J]. 吉林农业, 2010(6): 44-45, 49
- [12] Thurber RV, Haynes M, Breitbart M, et al. Laboratory procedures to generate viral metagenomes[J]. Nature Protocols, 2009, 4(4): 470-483
- [13] Hurwitz BL, Deng L, Poulos BT, et al. Evaluation of methods to concentrate and purify ocean virus communities through comparative, replicated metagenomics[J]. Environmental Microbiology, 2013, 15(5): 1428-1440
- [14] Nooij S, Schmitz D, Vennema H, et al. Overview of virus metagenomic classification methods and their biological applications[J]. Frontiers in Microbiology, 2018, 9: 749
- [15] Shi M, Lin XD, Chen X, et al. The evolutionary history of vertebrate RNA viruses[J]. Nature, 2018, 556(7700): 197-202
- [16] He JZ, Yuan CL, Shen JP, et al. Methods for and progress in research on soil metagenomics[J]. Acta Pedologica Sinica, 2012, 49(1): 155-164 (in Chinese)  
贺纪正, 袁超磊, 沈菊培, 等. 土壤宏基因组学研究方法与进展[J]. 土壤学报, 2012, 49(1): 155-164
- [17] Niu XG, Han M, Han XR. Metagenomics: new strategy for soil microbiology research[J]. Microbiology China, 2007, 34(3): 576-579 (in Chinese)  
钮旭光, 韩梅, 韩晓日. 宏基因组学: 土壤微生物研究的新策略[J]. 微生物学通报, 2007, 34(3): 576-579
- [18] Yegnasubramanian S. Chapter fifteen-preparation of fragment libraries for next-generation sequencing on the applied biosystems SOLiD platform[J]. Methods in Enzymology, 2013, 529: 185-200
- [19] Zhao HX, Gao YH, Ge Z, et al. Metagenomics in gastrointestinal microorganisms of animals[J]. Progress in Veterinary Medicine, 2013, 34(11): 106-109 (in Chinese)  
赵晗旭, 高云航, 葛铮, 等. 动物胃肠道微生物宏基因组学研究进展[J]. 动物医学进展, 2013, 34(11): 106-109
- [20] Schloss PD, Handelsman J. Biotechnological prospects from metagenomics[J]. Current Opinion in Biotechnology, 2003, 14(3): 303-310
- [21] Daniel R. The soil metagenome — a rich resource for the discovery of novel natural products[J]. Current Opinion in Biotechnology, 2004, 15(3): 199-204
- [22] Venter JC, Remington K, Heidelberg JF, et al. Environmental genome shotgun sequencing of the Sargasso Sea[J]. Science, 2004, 304(5667): 66-74
- [23] Paez-Espino D, Eloe-Fadrosh EA, Pavlopoulos GA, et al. Uncovering Earth's virome[J]. Nature, 2016, 536(7617): 425-430
- [24] Adriaenssens EM, van Zyl L, de Maayer P, et al. Metagenomic analysis of the viral community in Namib Desert hypoliths[J]. Environmental Microbiology, 2015, 17(2): 480-495
- [25] Adriaenssens EM, Kramer R, van Goethem MW, et al. Environmental drivers of viral community composition in Antarctic soils identified by viromics[J]. Microbiome, 2017, 5(1): 83
- [26] Ji XL. Bacteriophage diversity, abundance and its role in the production of dissolved organic carbon in Napahai Plateau Wetland[D]. Kunming: Doctoral Dissertation of Kunming University of Science and Technology, 2015 (in Chinese)  
季秀玲. 高原湿地纳帕海噬菌体的多样性、丰度及其与 DOC 关系初步研究[D]. 昆明: 昆明理工大学博士学位论文, 2015
- [27] Zhang ZY. Coevolution between *Pseudomonas fluorescens* W-6 and its cold-active Siphoviridae phage VW-6S[D]. Kunming: Master's Thesis of Kunming University of Science and Technology, 2018 (in Chinese)  
张中耀. 荧光假单胞菌 W-6 与其低温长尾噬菌体 VW-6S 共进化研究[D]. 昆明: 昆明理工大学硕士学位论文, 2018
- [28] Wang S. Infection/anti-infective gene changes of *Pseudomonas fluorescens* W-6 and its phage VW-6S in coecoluation[D]. Kunming: Master's Thesis of Kunming University of Science and Technology, 2019 (in Chinese)  
王爽. 荧光假单胞菌及其噬菌体共进化过程中感染/抗感染基因变化研究[D]. 昆明: 昆明理工大学硕士学位论文, 2019
- [29] Nishimura Y, Watai H, Honda T, et al. Environmental viral genomes shed new light on virus-host interactions in the ocean[J]. mSphere, 2017, 2(2): e00359-16
- [30] Brum JR, Ignacio-Espinoza JC, Roux S, et al. Patterns and ecological drivers of ocean viral communities[J]. Science, 2015, 348(6237): 1261498
- [31] Cai LL, Jørgensen BB, Suttle CA, et al. Active and diverse viruses persist in the deep sub-seafloor sediments over thousands of years[J]. The ISME Journal, 2019, 13(7): 1857-1864
- [32] Ng TFF, Kondov NO, Deng XT, et al. A metagenomics and case-control study to identify viruses associated with bovine respiratory disease[J]. Journal of Virology, 2015, 89(10): 5340-5349
- [33] Klima CL, Holman DB, Ralston BJ, et al. Lower respiratory tract microbiome and resistome of bovine respiratory disease mortalities[J]. Microbial Ecology, 2019, 78(2): 446-456

- [34] Sachsenröder J, Braun A, Machnowska P, et al. Metagenomic identification of novel enteric viruses in urban wild rats and genome characterization of a group A rotavirus[J]. *Journal of General Virology*, 2014, 95(12): 2734-2747
- [35] Fawaz M, Vijayakumar P, Mishra A, et al. Duck gut viral metagenome analysis captures snapshot of viral diversity[J]. *Gut Pathogens*, 2016, 8(1): 30
- [36] Young JC, Chehoud C, Bittinger K, et al. Viral metagenomics reveal blooms of anelloviruses in the respiratory tract of lung transplant recipients[J]. *American Journal of Transplantation*, 2015, 15(1): 200-209
- [37] Arkhipova K, Skvortsov T, Quinn JP, et al. Temporal dynamics of uncultured viruses: a new dimension in viral diversity[J]. *The ISME Journal Multidisciplinary Journal of Microbial Ecology*, 2018, 12(1): 199-211
- [38] Shkoporov AN, Hill C. Bacteriophages of the human gut: the “known unknown” of the microbiome[J]. *Cell Host & Microbe*, 2019, 25(2): 195-209
- [39] Moreno-Gallego JL, Chou SP, di Rienzi SC, et al. Virome diversity correlates with intestinal microbiome diversity in adult monozygotic twins[J]. *Cell Host & Microbe*, 2019, 25(2): 261-272.e5

(上接 p.2435)

## 征 稿 简 则

3.5 参考文献: 参考文献按文内引用的先后顺序排序编码, 未公开发表的资料请勿引用。我刊参考文献需要注明著者(文献作者不超过 10 人时全部列出, 多于 10 人时列出前 10 人, 后加“等”或“et al.”, 作者姓前名后, 名字之间用逗号隔开)、文献名、刊名、年卷期及页码。国外期刊名必须写完整, 不用缩写, 不用斜体。参考文献数量不限。

参考文献格式举例:

- [1] Mahajan-Miklos S, Tan MW, Rahme LG, Ausubel FM. Molecular mechanisms of bacterial virulence elucidated using a *Pseudomonas aeruginosa*-*Caenorhabditis elegans* pathogenesis model[J]. *Cell*, 1999, 96(1): 47-56
- [2] Cheng XY, Liu WW, Xu Y, Zhou NY. Screening and characterization of culturable hydrocarbon-degrading strains from the South and East China Seas[J]. *Microbiology China*, 2019, 46(5): 975-985 (in Chinese)  
程晓宇, 刘伟伟, 许楹, 周宁一. 中国东海和南海海域可培养烃类降解细菌的筛选及功能[J]. *微生物学通报*, 2019, 46(5): 975-985
- [3] Shen T, Wang JY. *Biochemistry*[M]. Beijing: Higher Education Press, 1990: 87 (in Chinese)  
沈同, 王镜岩. *生物化学*[M]. 北京: 高等教育出版社, 1990: 87
- [4] Liu X. Diversity and temporal-spatial variability of sediment bacterial communities in Jiaozhou Bay[D]. Qingdao: Doctoral Dissertation of Institute of Oceanology, Chinese Academy of Sciences, 2010 (in Chinese)  
刘欣. 胶州湾沉积物细菌多样性及菌群时空分布规律[D]. 青岛: 中国科学院海洋研究所博士学位论文, 2010

## 4 特别说明

4.1 关于测序类论文: 凡涉及测定 DNA 或氨基酸序列的论文, 请先通过国际基因库 EMBL (欧洲)或 GenBank (美国)或 DDBJ (日本), 申请得到国际基因库登录号 (Accession No.)后再投来。

4.2 关于版权: (1) 本刊只接受作者独立创作的原创性作品, 享有自主知识产权, 无抄袭问题; 文中相关内容不曾以各种语种在国内外公开发表过, 并且不存在学术伪造、一稿多投、同一学术成果多篇发表等问题; 论文不涉及泄密及其他与著作权有关的侵权问题; 全部数据真实可靠, 且数据、图表未曾正式发表。若来稿被发现存在上述问题, 编辑部调查核实后可随时终止流程, 已发表的将发布公告公开撤销发表, 并将作者列入黑名单, 本刊不再受理该作者任何稿件。作者文责自负。(2) 凡在本刊通过审稿、同意刊出的文章, 所有形式(即各种文字、各种介质)的版权均属本刊编辑部所有。作者如有异议, 敬请事先声明。(3) 对录用的稿件编辑部有权进行文字加工, 但如涉及内容的大量改动, 将请作者过目同意。

4.3 审稿程序及提前发表: (1) 来稿刊登与否由编委会最后审定。对不录用的稿件, 一般在收稿 2 个月之内通过 E-mail 说明原因, 作者登录我刊系统或关注绑定微信也可查看。稿件经过内审、初审、终审通过后, 作者根据编辑部返回的退修意见进行修改补充后上传修改稿, 编辑部复审通过后将发出稿件录用通知单, 稿件按照投稿先后排队发表。(2) 本刊对投稿的个人和单位一视同仁。坚持文稿质量为唯一标准, 对稿件采取择优先登的原则。

## 5 发表费及稿费

论文一经录用, 将在发表前根据版面收取一定的发表费并酌付稿酬、赠送样刊。

## 6 联系方式

地址: 北京市朝阳区北辰西路 1 号院 3 号中国科学院微生物研究所《微生物学通报》编辑部(100101)

Tel: 010-64807511; E-mail: tongbao@im.ac.cn; 网址: <http://journals.im.ac.cn/wwxtbcn>