



# 反刍动物瘤胃噬菌体的宏基因组学研究方法及进展

吴祎程<sup>1,2</sup>, 冉涛<sup>3\*</sup>, 周传社<sup>1,2\*</sup>, 谭支良<sup>1,2</sup>

(1. 中国科学院亚热带农业生态研究所 亚热带农业生态过程重点实验室 畜禽养殖污染控制与资源化技术国家工程实验室 湖南省动物营养与生理代谢实验室 农业部中南动物营养与饲料科学观测站,长沙 410125; 2. 中国科学院大学,北京 100049;3. 兰州大学草地农业科技学院,兰州 730020)

**摘要:** 病毒(virus)存在于各种动物胃肠道中,反刍动物瘤胃也不例外。瘤胃是一个复杂且多样的微生态系统,包括细菌、真菌、原虫、古生菌和病毒等多种微生物。目前,多组学技术和生物信息学分析手段被广泛用于瘤胃细菌群落的研究。有研究表明,饲料、环境及宿主基因型会对瘤胃细菌多样性和菌群结构产生影响;瘤胃细菌也会反过来影响宿主机体代谢、生长发育和健康状况。与瘤胃细菌相比,瘤胃病毒受到的关注相对较少。随着分子生物学技术的发展,高通量测序技术为病毒组学的研究提供了必要的支持。研究者们将宏基因组学运用于瘤胃病毒,在瘤胃液中发现了许多新型病毒,极大地丰富了人们对瘤胃病毒多样性的认知。本文从瘤胃噬菌体概念出发,重点阐述了瘤胃病毒 DNA 提取、测序及生物信息学分析手段,并结合其他环境病毒的研究方法进行探讨,最后对瘤胃病毒宏基因组研究的应用前景进行了总结。

**关键词:** 瘤胃;病毒;噬菌体;宏基因组;高通量测序

中图分类号:S852.65

文献标志码:A

文章编号:0366-6964(2022)01-0020-12

## Rumen Bacteriophage in Ruminants: Metagenome Analytical Methods and Research Progress

WU Yicheng<sup>1,2</sup>, RAN Tao<sup>3\*</sup>, ZHOU Chuanshe<sup>1,2\*</sup>, TAN Zhiliang<sup>1,2</sup>

(1. Key Laboratory for Agro-Ecological Processes in Subtropical Region, National Engineering Laboratory for Pollution Control and Waste Utilization in Livestock and Poultry Production, Hunan Provincial Key Laboratory of Animal Nutrition Physiology and Metabolic Process, South Central Experimental Station of Animal Nutrition and Feed Science in the Ministry of Agriculture, Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha 410125, China; 2. University of the Chinese Academy of Sciences, Beijing 100049, China; 3. College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou 730020, China)

**Abstract:** Viruses are ubiquitous to the gastrointestinal tracts of various types of animals, including the rumen of ruminants. The rumen is a dense and taxonomically diverse consortium of bacteria, fungi, protozoa, archaea and virus. Nowadays, most of the researches by multi-omics and

收稿日期:2021-06-02

基金项目:国家自然科学基金(U20A2057;31772632);国家重点研发计划(2018YFD0501903;2017YFD0500500);兰州大学“双一流”建设人才项目(561120213)

作者简介:吴祎程(1997-),女,湖北武汉人,硕士生,主要从事瘤胃环境病毒提取及功能研究,E-mail: wuyicheng19@mails.ucas.ac.cn, Tel: 0731-84615204

\*通信作者:周传社,主要从事反刍家畜蛋白质营养调控与农牧复合生态系统研究,E-mail: zcs@isa.ac.cn;冉涛,主要从事传统动物营养学、分子营养学及胃肠道微生物组学研究,E-mail: rant@lzu.edu.cn

bioinformatics have focused on rumen bacteria. Studies have shown that diet, environment and host genotype have great impact on the diversity and community structure of rumen bacteria; meanwhile, rumen bacteria also affect the host's metabolism, growth performance and health status in turn. Compared with rumen bacteria, rumen phages have received relatively fewer attention. With the development of molecular biology technology, high-throughput sequencing has paved way for the study of virome. Many novel viruses in rumen have been discovered by viral metagenomics, which largely enriches the understanding of the diversity of rumen viruses. Starting from the concept of rumen phage, this review focused on the analytical methods including virus DNA extraction, sequencing and bioinformatics. Then discussed the virome analytical methods both in the rumen and other ecosystems, and finally summarized the application prospects of rumen virome research.

**Key words:** rumen; virus; bacteriophage; metagenomics; high-throughput sequencing

\* **Corresponding authors:** ZHOU Chuanshe, E-mail: zcs@isa.ac.cn; RAN Tao, E-mail: rant@lzu.edu.cn

病毒(virus)广泛存在于地球各种环境中,是地球上分布最广、数量最为巨大的生命实体<sup>[1]</sup>。研究发现,大部分病毒能特异性侵染细菌,这些病毒被称作噬菌体(bacteriophage 或 phage)<sup>[2-3]</sup>。在噬菌体被发现后不久,研究人员便证实了反刍动物瘤胃内也存在噬菌体<sup>[4-5]</sup>。长期以来,尽管研究人员意识到病毒可能在反刍动物瘤胃微生态系统中发挥着重要作用,但受限于技术手段和瘤胃环境的复杂性,瘤胃环境中噬菌体研究进展相对缓慢<sup>[6]</sup>。目前,对瘤胃噬菌体的认知多来自于传统的培养、显微镜观察和单一噬菌体的基因组学研究<sup>[7-8]</sup>,对其在瘤胃微生态系统层面的认识仍十分匮乏且片面,以致至今对其在瘤胃微生态系统中的组成和功能仍知之甚少。近年来,随着宏基因组学技术的发展,极大地推动了海洋生态环境中病毒生态学研究的发展,并取得了引人注目的成绩<sup>[9]</sup>。宏基因组学也被逐步运用于胃肠道病毒组(virome)的研究,同时刷新了研究人员对胃肠道噬菌体的复杂性和多样性的认知<sup>[10]</sup>。Marrique 等<sup>[11]</sup>首次利用病毒宏基因组学技术发现肠道中 81%~93%的噬菌体为新型噬菌体,由于病毒缺乏合适的标记基因和分析基准,它们无法被归类或找到对应宿主。瘤胃微生态环境复杂,瘤胃作为一个巨大的基因资源库,具有巨大的新病毒发掘空间,病毒宏基因组学技术与传统技术的联用在这一方面具有独特优势,进而推动瘤胃噬菌体相关研究。

本文主要聚焦反刍动物瘤胃生态系统中的噬菌体群落,并讨论瘤胃噬菌体研究方法及现状,及从其他生境中的噬菌体研究所得到的启发。本文还展望

了病毒宏基因组学在瘤胃环境中的应用前景,以期后续瘤胃噬菌体组研究提供科学参考。

## 1 瘤胃噬菌体概述

反刍家畜瘤胃内栖息着数以万亿计的微生物<sup>[12-15]</sup>(图 1),包括细菌、真菌、原虫、古生菌和病毒,其中病毒的数量( $>10^8 \sim 10^9$  pfu · mL<sup>-1</sup>)仅次于细菌,且主要为噬菌体(bacteriophage 或 phage)<sup>[16]</sup>。瘤胃内的噬菌体具有不同形态,现有研究表明,反刍动物瘤胃中,有尾噬菌体目(Caudovirales)普遍存在,包括肌尾噬菌体科(Myoviridae)、长尾噬菌体科(Siphoviridae)、短尾噬菌体科(Podpoviridae)占主导地位<sup>[17-18]</sup>,以及最近被鉴定出来的阿克曼噬菌体科(Ackermannviridae)和赫雷尔噬菌体科(Herelleviridae)<sup>[19]</sup>。其实,早在 20 世纪 60 年代,“瘤胃噬菌体”一词就已进入人们视野,但对其在瘤胃微生态系统中的功能知之甚少,直到 20 世纪 80 年代末,研究人员才开始对其进行探索<sup>[20]</sup>。随着测序技术的发展,宏病毒组学技术日臻完善,借助该技术手段能够更好地探索噬菌体这个“暗物质”。已有研究表明,噬菌体在塑造瘤胃菌群结构、维持微生物多样性和调节宿主菌代谢等方面发挥着重要作用<sup>[21-22]</sup>。对反刍动物瘤胃噬菌体多样性的研究表明,噬菌体可能是通过裂解宿主细菌从而影响瘤胃细菌种群动态变化<sup>[7,23-24]</sup>。Anderson 等<sup>[23]</sup>通过营养调控试验首次报道了饲料变化对反刍动物瘤胃噬菌体群落的影响,并表明,噬菌体可通过其编码的辅助代谢基因(auxiliary metabolic genes, AMGs),经

过裂解宿主细菌、产能、复制和微生物代谢的重新编程等一系列过程来影响瘤胃微生态系统<sup>[17,21,23]</sup>,而其中宿主菌代谢与反刍动物饲料转化效率密切相关。众所周知,瘤胃微生物群体结构的稳定性和多

样性对于反刍家畜的健康、营养、免疫和生存至关重要,因此,需要重视瘤胃噬菌体在整个瘤胃微生态系统中的地位、功能及作用。

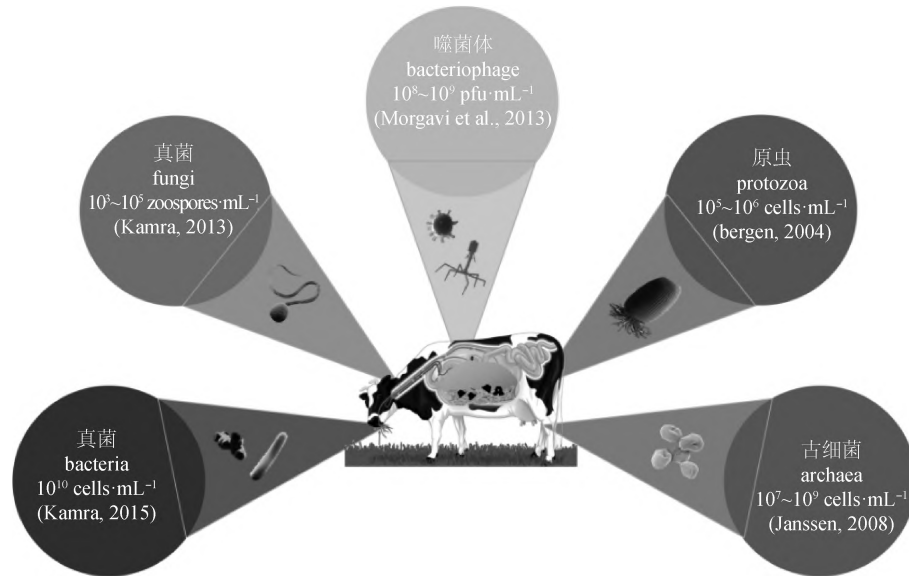


图1 瘤胃微生物群的主要分类及丰度<sup>[12-15]</sup>

Fig.1 Main taxonomic groups of the rumen microbiome and the domain level<sup>[12-15]</sup>

## 2 瘤胃噬菌体研究方法

研究噬菌体的传统方法是显微镜观察和体外分离培养,但在实验室环境中培养的宿主菌数量有限;噬菌体还具有形态小、遗传进化快等特点,缺乏细菌所拥有的相对保守的系统发育标记,如16S rRNA基因;且感染不同宿主细菌的噬菌体具有高度特异性,不同噬菌体序列的重叠群非常小<sup>[11,25-26]</sup>。在瘤胃噬菌体的早期研究中,利用透射电子显微镜(transmission electron microscope, TEM)技术,研究人员发现,瘤胃液存在着形态各异、高度多样化的噬菌体种群,并猜测其可能影响瘤胃内细菌的种群<sup>[27]</sup>。后来,随着DNA分析技术、脉冲场凝胶电泳技术(pulsed field gel electrophoresis, PFGE)和微生物分离鉴定技术的发展,研究人员证实,反刍动物瘤胃液中存在大量溶菌性噬菌体<sup>[7]</sup>,并陆续分离鉴定了多种,包括以坏死梭杆菌(*Fusobacterium necrophorum*)<sup>[28]</sup>、白色瘤胃球菌(*Ruminococcus albus*)<sup>[29]</sup>、溶纤维丁酸弧菌(*Butyrivibrio fibrisolvens*)<sup>[30]</sup>和瘤胃拟杆菌(*Prevotella ruminicola*)<sup>[31]</sup>等作为宿主菌的噬菌体。TEM技术也随之被广泛用于新型噬菌体分离株的形态学表征<sup>[25, 32]</sup>。如今,

宏基因组学测序技术和生物信息学的快速发展,使研究人员能更深入地了解病毒群落的结构和功能。宏基因组学(metagenomics)手段是以某一特定环境样本中的微生物群体基因组为研究对象,以功能基因筛选和序列测定分析为研究手段,以微生物多样性、种群结构、进化关系、功能活性、相互作用以及其与环境间的关系为研究目的的一种新的微生物研究方法<sup>[33]</sup>。而宏病毒组学(viral metagenomics)则是结合病毒自身特点,将宏基因组学方法应用于病毒学领域<sup>[34]</sup>。

瘤胃噬菌体宏基因组学研究,主要由病毒样颗粒(virus-like particles, VLPs)浓缩液制备与生物信息学分析两部分组成(图2)。具体步骤包括:1)瘤胃样本采集,经过滤、消化以及高速离心等技术,去除细菌等其他潜在宿主;2)对样品进行富集、纯化,得到病毒浓缩液;3)根据病毒核酸类型,选择合适的病毒核酸提取试剂盒或手工提取的方式获取病毒核酸,构建病毒基因组文库,并通过高通量测序技术进行宏病毒组测序;4)对测序所得原始数据进行质控,基于重叠区(overlap)将高质量测序读段(reads)拼接为重叠群(contigs);5)通过病毒序列鉴定软件在重叠群中鉴别、筛选病毒序列,从而了解该

环境中病毒构成;6)对病毒基因组开放阅读框(open reading frame, ORF)进行预测,再通过注释

工具将 ORF 与多个数据库(KEGG、COG、CAZy、CARD 等)比对进行功能基因注释。

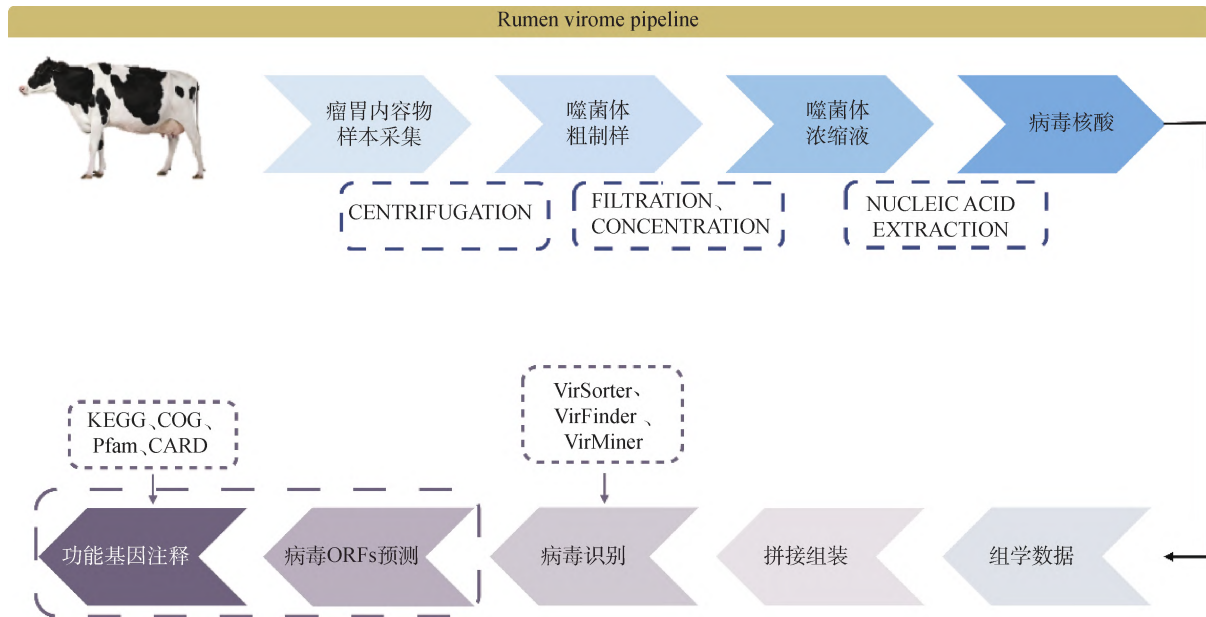


图 2 瘤胃病毒组分析流程图

Fig. 2 Flowcharts of rumen virome analysis

## 2.1 样品前处理及噬菌体核酸提取

瘤胃噬菌体即为那些以瘤胃细菌为宿主菌的噬菌体,或是那些存在于反刍动物瘤胃体内的噬菌体<sup>[31]</sup>。由于采样后,细菌和噬菌体仍然彼此接触,噬菌体的侵染作用仍持续存在,长时间孵育会影响噬菌体与微生物比例<sup>[35]</sup>。因此,样品采集后要立即超低温冻存或立即对样品进行处理,将病毒样颗粒从样品中分离并纯化,用于病毒遗传物质提取。现有的提取瘤胃噬菌体核酸的方法大多参考人或小鼠肠道噬菌体及土壤噬菌体分离方法<sup>[17,23,31]</sup>,截至目前,仅有一篇文献以山羊和绵羊瘤胃液作为瘤胃噬菌体来源,建立了瘤胃噬菌体富集及 DNA 提取方法(包括过滤、离心、沉淀及核酸提取一系列过程)<sup>[18]</sup>。

瘤胃环境异常复杂且微生物多样性高,瘤胃样品中噬菌体的分离质量直接关系到病毒宏基因组数据的质量。Hungate<sup>[36]</sup>研究表明,大约 75% 的瘤胃细菌紧密附着在饲料颗粒上,以细菌为宿主的噬菌体也会附着在饲料上,因此,研究瘤胃噬菌体时,需同时采集固相和液相样品,以便充分囊括瘤胃内的所有噬菌体。在从样品中分离噬菌体时,常利用噬菌体与其他微生物颗粒大小不同这一特征,借助微孔滤膜、离心等方式来实现噬菌体与其他微生物的分离。但是,实际操作中噬菌体样极易混入细菌、真

菌等微生物,尤其是那些与噬菌体颗粒大小相近的微生物,使得提取的病毒宏基因组中常混有细菌、真菌等其他微生物的核酸<sup>[37]</sup>。于是,在制备噬菌体粗制样时,微孔滤膜、化学试剂等的选择对去除细菌等其他微生物的污染和保证噬菌体群体的代表性具有重要影响,进而对后续病毒核酸纯度、基因功能注释的准确性产生影响。Conceição-Neto 等<sup>[38]</sup>研究发现,常见代表性病毒的大小由 17 到 1 000 nm 不等,其中,常见噬菌体的大小为 125 nm 左右。胃肠道病毒颗粒中绝大部分是噬菌体,为了避免细菌的污染,在研究胃肠道噬菌体组时,常选择孔径为 0.22 或 0.45  $\mu\text{m}$  的微孔滤膜,当样本环境复杂时,为避免遗漏巨型病毒信息,常选择 0.45  $\mu\text{m}$  的微孔滤膜,但为了更有效地防止噬菌体宿主菌污染,0.22  $\mu\text{m}$  的微孔滤膜仍被广泛结合使用<sup>[39]</sup>。Friedersdorff 等<sup>[30]</sup>选择 0.45 和 0.22  $\mu\text{m}$  两种孔径的微孔滤膜依次过滤瘤胃液样品,最终分离得到 5 种瘤胃溶菌性噬菌体,并对其进行了全基因组学测序。同样,Namonyo 等<sup>[18]</sup>在建立瘤胃病毒 DNA 提取方法时也选用了 0.45 和 0.22  $\mu\text{m}$  两种孔径的微孔滤膜依次进行过滤。而 Klieve 等<sup>[31]</sup>只选择了孔径为 0.45  $\mu\text{m}$  的微孔滤膜对瘤胃相关样本进行过滤。虽然现有文献表明暂未鉴定出巨型瘤胃噬菌体,但并不排除技术手

段造成的误差,为了更全面地研究瘤胃环境噬菌体群落,研究人员仍需根据研究目的谨慎选择微孔滤膜的孔径。

在获得含噬菌体的粗制样后,常利用聚乙二醇(polyethylene glycol, PEG)配合不同浓度的 NaCl 来沉淀噬菌体颗粒以实现噬菌体的浓缩。较为常用的 PEG 和 NaCl 用量为 25% PEG6000( $w/v$ )和  $1.0 \text{ mol} \cdot \text{L}^{-1}$  NaCl<sup>[18]</sup>; 10% PEG8000 和  $0.5 \text{ mol} \cdot \text{L}^{-1}$  NaCl<sup>[31]</sup> 及 20% PEG6000 和  $2.5 \text{ mol} \cdot \text{L}^{-1}$  NaCl<sup>[30]</sup>。添加 PEG 后,常需在低温下(4 °C)孵育过夜,以使噬菌体充分沉淀,经过高速离心( $13\ 000 \times g$ , 30 min, 4 °C;  $12\ 000 \times g$ , 30 min, 4 °C 或  $52\ 350 \times g$ , 10 min, 4 °C)后,倒掉上清液,获得的沉淀即为噬菌体样品。根据试验需求,该样品可以直接用于后续试验,也可经超速离心进一步纯化后用于后续试验。较为常用的是氯化铯(CsCl)密度梯度离心等,此法可对特定密度范围内的噬菌体进行纯化,并根据密度将 VLPs 分层。Carroll-Portillo 等<sup>[40]</sup>的测序结果表明,用氯化铯密度梯度离心法富集的 VLPs 样品宿主细菌核酸去除率比未经富集的高,但氯化铯密度梯度离心法在富集样品时,对不同密度的噬菌体有明显的偏好性。Cordova 等<sup>[41]</sup>也表明,某些噬菌体会由于氯化铯产生的渗透压胁迫(osmotic shock)而导致噬菌体核酸丢失,因此,研究人员会选择性采用氯化铯密度梯度离心法。

与瘤胃微生物组研究类似,病毒组研究也需先从分离得到的样本中提取遗传物质。鉴于 DNA 病毒在瘤胃环境中的优势地位,目前,对于瘤胃噬菌体的研究主要集中于 DNA 病毒的核酸提取<sup>[20]</sup>。噬菌体核酸提取主要有两种方式:第一种是试剂盒法,如 Roche high pure viral nucleic acid large volume kit (Roche, Switzerland)<sup>[18]</sup>、QIAamp Ultra Sens Virus Kit (Qiagen)<sup>[23]</sup> 和 Fast DNA<sup>TM</sup> Spin Kit for Soil (MP Biomedicals, Solon, OH, United States)<sup>[31]</sup> 等,这类试剂盒可用于 DNA 病毒核酸的提取,试剂盒具有操作方便、提取纯度高且不需要直接接触有害化学试剂等优点;第二种是传统的苯酚-氯仿提取法,利用核酸、蛋白质等杂质在水相和有机相中溶解性不同而重新分配的性质来实现核酸提取,该方法价格低廉,但耗时长且提取浓度低,常需进一步纯化后才能达到测序要求<sup>[42]</sup>。试验中,若遇提取的噬菌体 DNA 浓度过低而不能满足测序要求的情况,可利用多重置换扩增(multiple displace-

ment amplification, MDA)对核酸进行扩增,得到的 DNA 可用于构建基因文库或高通量测序<sup>[43]</sup>,但扩增技术所得到的产物可能存在偏好性,还是建议提高原始噬菌体样品富集量以保证结果的准确性<sup>[3]</sup>。

## 2.2 高通量测序与病毒序列鉴定

高通量测序技术(high-throughput sequencing, HTS)是对传统 Sanger 测序技术革命性的变革,可对数百万个 DNA 分子进行同时测序,并可深入地对一个物种的基因组和转录组进行整体分析,因此,也称其为下一代测序技术(next generation sequencing, NGS)<sup>[44]</sup>。HTS 具有通量高、速度快、成本低等优点,已广泛应用于宏病毒组学的研究中。近年来,牛津纳米孔技术(Oxford nanopore technology, ONT)作为新兴的单分子实时测序技术(single molecule real-time sequencing technology, SMRT)之一,具有快速制备文库,超长读取和实时数据采集等优势<sup>[45]</sup>。该技术的核心是蛋白质纳米孔,通过产生覆盖单个病毒颗粒内所有突变的基因组长度读数来获得病毒基因组。在此基础上,中国科学院微生物研究所王军研究员团队开发了一种新的工作流程(包括病毒颗粒富集、核酸的逆转录和扩增以及生物信息学分析),并首次使用 ONT PromethION 平台对人类病毒组进行表征<sup>[46]</sup>。相较于传统的二代测序,ONT 具有读长长的优点,因此,无需进行 PCR 扩增,便可以在更短时间内、更准确地从样品中生成数据。

高通量测序后首先要对原始数据进行质量控制和过滤,去除宿主核酸,再进行序列的拼接,并运用专业软件来鉴定病毒序列,确定该序列的生物群落来源,筛选出有用的基因信息。对噬菌体基因组进行组装时,一般挑选 1 000 bp 以上组装序列,再应用一系列软件对病毒 contig 进行鉴定。表 1 列出了常用于噬菌体鉴定的生物信息学分析工具。

从同时包含噬菌体及其宿主的混合基因组数据集中准确识别噬菌体,是分析样品中噬菌体组成的关键步骤。目前,有非常多基于序列或基于重叠群对病毒进行鉴定的生物信息学分析软件。VirSorter (<https://github.com/simroux/VirSorter>)可利用概率模型和大量病毒数据最大限度地检测新病毒,虽然该工具在很大程度上依赖对已有病毒基因组的相似性搜索,但它有一个显著优势,即可手动管理病毒参考基因组数据库,并补充了以海水、人体肠道、肺部组织及唾液样本为来源的宏病毒组序列<sup>[47]</sup>。2017年,Ren 等<sup>[48]</sup>开发的 VirFinder (<https://github.com>).

表 1 病毒鉴定工具

Table 1 Tools for the identification of viruses

分析工具 Software	特点 Characteristic	发表杂志 Journal	发表年份 Year	被引量 Citation	参考文献 Reference
VirSorter	可手动管理病毒参考基因组数据库	<i>PeerJ</i>	2015	325	[47]
VirFinder	对于大小 < 1 000 bp 的病毒 contig 鉴定效果较好	<i>Microbiome</i>	2017	120	[48]
MARVEL	可对 Caudovirales 目 dsDNA 病毒有效注释	<i>Frontiers in Genetics</i>	2018	35	[49]
VirMiner	可预测噬菌体-宿主相互作用	<i>Microbiome</i>	2019	16	[50]
VIBRANT	可预测到其他分析工具无法预测到的 prophage	<i>Microbiome</i>	2020	14	[51]
VirSorter2	能一定程度地减少真核基因组和质粒的错误识别	<i>Microbiome</i>	2021	0	[52]

com/jessieren/VirFinder)是通过利用细菌和病毒在 K-mer 上的差异,将病毒从宏基因组数据集中识别出来,与 VirSorter 相比,它能鉴定出更多潜在病毒,尤其是大小在 1 000 bp 以下的病毒 contig。但 VirSorter 和 VirFinder 进行噬菌体分析的功能仍相对有限,在鉴定噬菌体重叠群后,无法预测噬菌体与宿主间的相互作用。MARVEL(<https://github.com/LaboratorioBioinformatica/MARVEL>) 是 2018 年由 Amgarten 等<sup>[49]</sup>开发的工具,以其对 Caudovirales 目 dsDNA 病毒的有效注释能力而为人所知,但该工具鉴定 RNA 病毒的错报率非常高。2019 年开发的 VirMiner (<https://github.com/TingtZHENG/VirMiner>)则针对噬菌体分析功能做出了改进,可通过预测模型从宏基因组数据集中预测噬菌体重叠群并进行下游分析,包括功能基因注释及噬菌体与宿主间关系预测等<sup>[50]</sup>。VIBRANT (<https://github.com/AnantharamanLab/VIBRANT>),是 2020 年开发的利用迭代注释进行病毒识别的工具,它能高效识别以细菌为宿主的多种病毒(包括 dsDNA、ssDNA、dsRNA 和 ssRNA 病毒),并从宿主序列中提取前病毒区域,从而预测到其他分析工具无法预测到的前病毒(prophage),且能解析环境中不同病毒之间新陈代谢的能力<sup>[51]</sup>。2021 年,Guo 等<sup>[52]</sup>改进并发布了 VirSorter2 (<https://github.com/jiarong/VirSorter2>),但 VirSorter 仍是最被广泛使用的工具。现有的每种工具都有其局限性,应根据实际需求选择使用或联合使

用不同的软件。

### 2.3 病毒功能基因注释

使用 BLAST 软件进行基因预测,并将获得的非冗余基因组与 RefSeq 病毒数据库的参考序列进行比对,获取该基因的功能信息<sup>[53]</sup>。通过对病毒功能基因注释,能够对深入认识病毒个体生命过程提供理论基础,还有助于了解病毒群落的生态过程及与宿主群落的互作网络,从而阐释病毒与宿主间复杂的相互作用机制。

京都基因与基因组百科全书(Kyoto encyclopedia of genes and genomes,KEGG)是一个系统分析基因功能的知识库,KEGG 具有强大的图形功能,可利用图形来介绍众多代谢途径(包括各种代谢通路、合成通路、膜转运、信号传递、细胞周期以及疾病相关通路等)以及各个途径间的关系,这样可以直观地反映基因与相关代谢的关系<sup>[54]</sup>。蛋白质直系同源簇(clusters of orthologous groups of proteins, COG)数据库,可将不同物种中的直系同源基因进行聚类,通过在不同物种中建立相关同源蛋白簇来预知位置蛋白质的功能,同时也为分子系统发育分析提供数据基础。2015 年,Galperin 等<sup>[55]</sup>基于完整噬菌体基因组中的编码蛋白系统进化关系构建了 POG(phage orthologous groups)数据库,可在双链 DNA 噬菌体的全基因组序列中鉴定出保守的直系同源簇。Pfam 数据库是蛋白质家族的数据库,根据多序列比对结果和隐马尔可夫模型(hidden markov model, HMMs),可查询蛋白质家族或蛋白结构域

的注释、结构及多序列比对信息,被广泛用于基因功能注释<sup>[56]</sup>。抗生素的滥用对畜牧业造成了深远影响,虽然抗生素对病毒并无直接作用,但有研究发现,瘤胃噬菌体携带抗生素耐药性基因<sup>[23]</sup>,因此,对病毒功能基因序列进行耐药性检测就显得尤为重要。抗生素综合研究数据库(comprehensive antibiotic resistance database, CARD)可在细菌耐药性的分子基础上,提供参考 DNA 和蛋白质序列、检测模型和生物信息学工具,通过与该数据库进行比对,可用于关联抗生素模块及其目标、抗性机制、基因变异等信息与耐药基因相关的注释信息<sup>[57]</sup>。

### 3 瘤胃噬菌体基因组研究趋势及进展

借助 Web of Science 核心合集以“rumen bacteriophage”、“rumen phage”或“rumen virus”为关键词进行检索,2000—2021 年时间段内共有 108 篇相关文献,且发文量呈现整体上升趋势(图 3a),2021 年发表的 5 篇文献并未在图中标明。其中,美国、澳大利亚、印度、新西兰和加拿大的发文量分别为 33.66%、15.39%、9.62%、8.65% 和 7.69%,而我国在瘤胃噬菌体这一领域的研究严重滞后。所发表的关于瘤胃噬菌体的文章中,基于宏基因组学的文献仅有 7 篇<sup>[17-18,21,23-24,58-59]</sup>,涉及研究动物包括水牛、奶牛、绵羊、山羊和麋鹿,且很少涉及瘤胃噬菌体群落动态变化规律及瘤胃噬菌体在整个瘤胃生态系统中的作用。随后,利用 VOSviewer 进行关键词密度聚类,并进行可视化分析<sup>[60]</sup>(图 3b),图中结点越接近黄色,表示此为研究热点,可发现瘤胃环境病毒研究热点为噬菌体及宏基因组层面。而以“gut phage”或“gut virus”作为关键词检索,在 2000—2021 年时间段内共有 4 449 篇相关文献,这说明,瘤胃噬菌体是一个极具发掘空间的研究领域。

对病毒进行全基因组测序(complete genome sequencing),又叫单病毒基因组学(single-virus genomics, SVG),可获得特定病毒(噬菌体)的全部核酸序列。SVG 即以纯培养的病毒为研究对象,进行核酸提取、文库构建并对单个病毒进行全基因组测序,SVG 从头组装可以规避病毒遗传物质多样性的问题,从而构建更完整的病毒基因组,以获得更全面的病毒数据库的信息<sup>[61]</sup>。目前,仅有两篇基于 SVG 来研究瘤胃噬菌体的文章,表 2 综述了噬菌体及其宿主信息。

Friedersdorff 等<sup>[30]</sup>以瘤胃拟杆菌(*Prevotella*

*ruminicola*)、白色瘤胃球菌(*Streptococcus bovis*)及瘤胃链球菌(*Ruminococcus albus*)为宿主,采用双层平板法(plaque assay)从瘤胃相关样品(废水、牛粪及瘤胃液)中分离出 5 株烈性噬菌体,这是首次对瘤胃环境中以特定细菌为宿主的噬菌体的基因组研究。随后对这 5 株噬菌体进行全基因组测序,并通过 Glimmer 和 Prodigal 在线工具对噬菌体全基因组序列进行潜在编码序列(coding sequence, CDS)预测和功能注释。发现在  $\phi$ Brb01 和  $\phi$ Brb02 基因组中,CDS 可分为 DNA 复制转录(DNA replication and transcription)、DNA 包装(DNA packaging)及宿主裂解(Host lysis)3 个功能组。溶纤维丁酸弧菌(*Butyrivibrio fibrisolvens*)是瘤胃内的一种优势菌,对纤维降解和蛋白质分解等有重要作用。Friedersdorff 等<sup>[30]</sup>以其为宿主,从瘤胃液及粪便中分离了 5 株噬菌体,并进行全基因组测序和利用 Glimmer 和 GeneMarkS 在线工具进行功能注释。结果显示,在科水平上,瘤胃噬菌体以有尾噬菌体为主,其中长尾噬菌体科占主导地位,与前人研究结果一致<sup>[17-18]</sup>。*Ceridwen* 属噬菌体基因组中与其他噬菌体基因组同源的 ORF 中,其中 20 个与 Siphoviridae 科噬菌体有同源性,1 个与 Podoviridae 科噬菌体有同源性。另外,用 PHACTS 软件对噬菌体生活史进行预测发现,这 10 株噬菌体中仅有 *Arawn* 和 *Idris* 属噬菌体基因组中存在编码整合酶的基因,表明这两株噬菌体可能为溶源性噬菌体(lysogenic phage),其他 8 株噬菌体为溶菌性噬菌体(lytic phage)。值得注意的一点是,从瘤胃液中分离出的 Bo-Finn 噬菌体与从牛粪便中分离出的 *Arian* 噬菌体基因组序列相似性可达 98.6%,它们属于同种噬菌体。

### 4 不同生境噬菌体研究进展

自 2002 年美国学者 Breitbart 等<sup>[62]</sup>首次对海洋病毒群(virome)进行宏基因组测序,到 2020 年,全球范围内的研究已报道了近 20 万个病毒种群<sup>[63]</sup>。目前,噬菌体的研究已然成为研究不同生境内微生物群落结构和功能的焦点,如哺乳动物肠道噬菌体与机体健康和疾病发生已受到广泛关注。关于不同生境噬菌体(如人体肠道、土壤等)的研究已证实,噬菌体是导致环境中宿主细菌死亡的主要原因<sup>[64]</sup>,且噬菌体可作为水平转移基因(horizontal gene transfer, HGT)的重要载体,与宿主菌进行遗传物质的

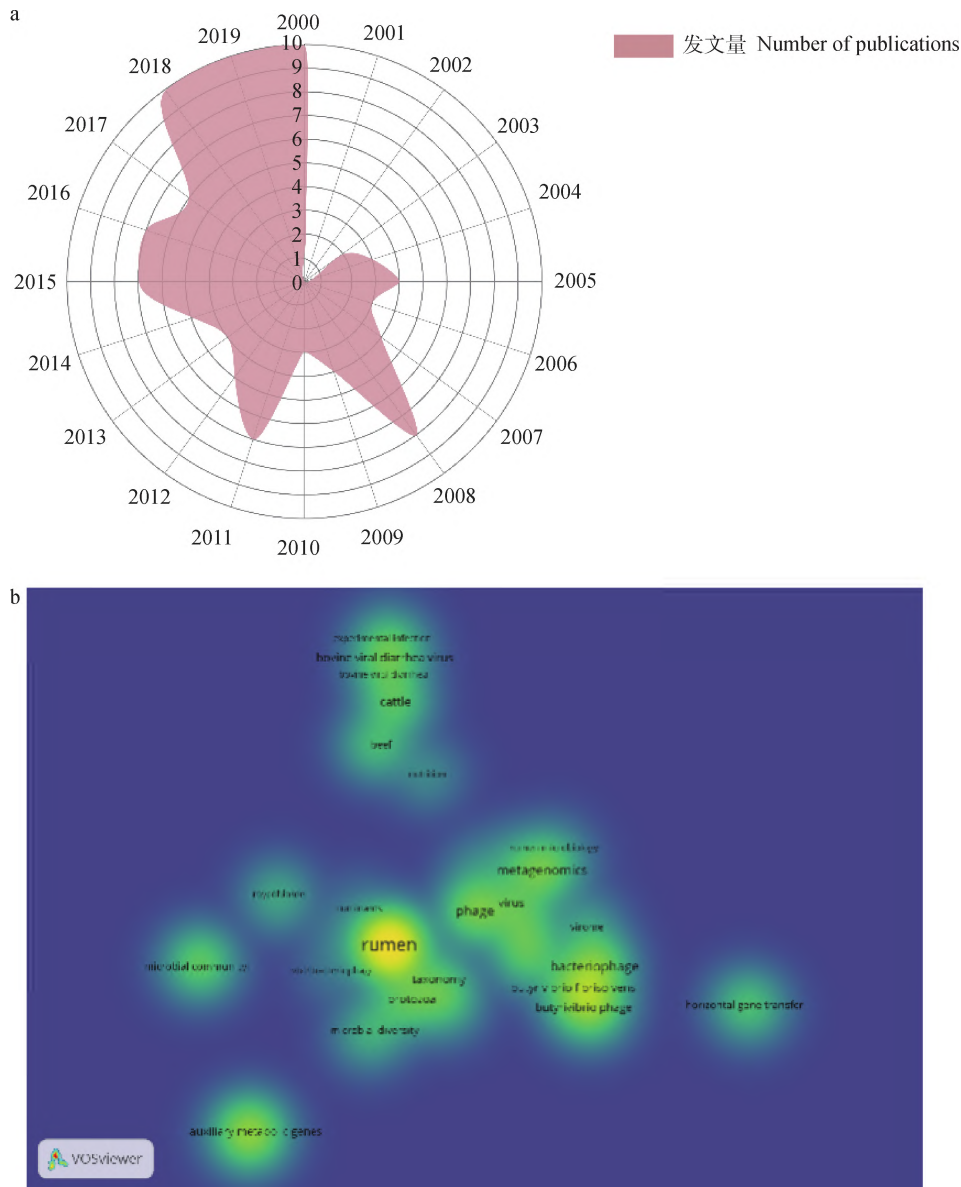


图3 瘤胃噬菌体领域年发表论文数量(a)及关键词密度图谱(b)

Fig. 3 Number of publications(a) and keyword density network map(b) in the field of rumen virus (bacteriophage)

交换,进而调节细菌的进化和多样性;值得注意的是,HGT也是导致抗生素抗性基因(antibiotic resistance genes, ARGs)扩散的重要因素<sup>[65]</sup>。鉴于瘤胃内噬菌体研究相对滞后的现状,可以适度借鉴其他环境样本中的噬菌体研究方法并用于瘤胃噬菌体研究。

许多环境噬菌体的研究得益于微生物多组学方法联用<sup>[66-67]</sup>,取得了喜人的成果。例如,Emerson等<sup>[66]</sup>利用宏基因组和宏转录组揭示了气候变化中噬菌体对复杂碳降解的潜能。此外,Brum等<sup>[67]</sup>结合蛋白质组学和宏基因组学,从海洋环境中鉴定出多样性极高的病毒衣壳蛋白。类似的,蛋白质组学

这一新兴领域在瘤胃微生态系统中的研究表明,瘤胃噬菌体介导的细胞裂解会释放微生物胞内酶,包括那些参与碳水化合物分解的微生物酶,这些酶的释放会促进瘤胃内饲料的降解<sup>[59]</sup>。

除了以上能通过组学技术挖掘出的数据外,病毒宏基因组测序结果中大部分序列为未知序列,为了充分利用这些未知数据,Brum等<sup>[67]</sup>通过对预测的蛋白质序列进行全长相似性聚类,揭示了七大大洋海洋样本中存在一个核心病毒基因集,这样可以从未知序列中得到另一种“已知”结果。Swain等<sup>[22]</sup>考虑到瘤胃病毒在不同个体间存在的差异,运用类似方法发现了核心瘤胃病毒基因集。此外,



表 2 已报道的获得全基因组信息的瘤胃噬菌体

Table 2 Rumen phages with genome-wide information reported

噬菌体名	科水平分类	基因组大小/kb	预测宿主菌	总 ORFs/个
Phage	Family	Genome length	Predicted host	Total ORFs
φBrb01	Siphoviridae	33.91±2.41	<i>Prevotella ruminicola</i>	46
φBrb02	Siphoviridae	33.02±1.16	<i>Prevotella ruminicola</i>	48
φRa02	Podoviridae	12.80	<i>Ruminococcus albus</i>	16
φRa04	Podoviridae	14.00	<i>Ruminococcus albus</i>	15
φSb01	Siphoviridae	30.90±4.4	<i>Streptococcus bovis</i>	46
Arawn	Siphoviridae	31.12	<i>Butyrivibrio fisolvens</i>	50
Arian	Siphoviridae	33.50	<i>Butyrivibrio fisolvens</i>	54
Bo-Finn	Siphoviridae	33.23	<i>Butyrivibrio fisolvens</i>	51
Ceridwen	Siphoviridae	39.75	<i>Butyrivibrio fisolvens</i>	73
Idris	Siphoviridae	31.13	<i>Butyrivibrio fisolvens</i>	55

Allen 等<sup>[68]</sup>研究海洋微生物群落时采用了一种基于噬菌体基因组标签的方法 (phage genome signature-based recovery, PGSR), 从宏基因组中提取噬菌体序列, 并结合病毒/细菌比 (virus-to-bacteria ratio, VBR) 来评估病毒对细菌群落的影响, 该方法有助于研究人员了解噬菌体在塑造整个海洋微生物群落结构中的作用。但目前未见此法运用于瘤胃噬菌体研究。

值得注意的是, 不同生态环境中研究方法的适用性有限。即使两个生态系统具有非常相似的病毒群落结构, 他们潜在的微生态关系也不尽相同。比如, 极地水体环境噬菌体和胃肠道环境噬菌体中, 溶源性噬菌体较溶菌性噬菌体在数量上占有优势<sup>[69-71]</sup>。但极地海洋生态系统中, 随着细菌丰度增加, 溶源性噬菌体会从溶源循环 (lysogenic cycle) 转变为溶菌循环 (lytic cycle)<sup>[69]</sup>; 而在肠道中, 宿主菌丰度增加时, 溶源性噬菌体会持续占领主导地位<sup>[70-71]</sup>。因此, 在将不同生境中噬菌体研究方法 & 生物学概念外推到瘤胃环境时仍需谨慎。

## 5 瘤胃宏病毒组的前景与展望

随着宏基因组学和生物信息学的发展, 研究人员可以不再需要体外分离培养噬菌体, 就能获得一些分散、丰度低的病毒的结构和功能信息, 也逐步增进对瘤胃内噬菌体的多样性及潜在功能的认知。但是, 现阶段学术界对瘤胃噬菌体及其功能基因的科学认知仍然十分有限, 笔者结合自身体会, 认为今后

此领域的研究应主要聚焦在以下几个方面:

**5.1** 宏基因组学研究中, 瘤胃噬菌体样品制备时主要通过微孔滤膜过滤的方式进行纯化富集, 因此, 难免会因为滤膜孔径选择过小而造成颗粒较大的噬菌体的遗传信息的遗漏。现阶段瘤胃噬菌体的提取、宏病毒组分析等缺乏统一技术规范, 建立成熟的噬菌体分离纯化实验室标准化操作流程以及针对巨型噬菌体的优化提取和富集技术是当务之急。

**5.2** 目前的研究主要聚焦在瘤胃 DNA 病毒, 对于 RNA 病毒的研究较少。未来要利用多组学手段 (如宏基因组、宏蛋白组、宏转录组和宏代谢组等) 结合生物信息学技术, 同步关注瘤胃 RNA 病毒的基因功能, 尽可能获得全面的噬菌体信息, 这有助于探明瘤胃整体病毒的生态功能和作用机制。

**5.3** 鉴于噬菌体种群在个体间差异显著且噬菌体种群存在于反刍动物整个消化道, 需进一步研究不同生理阶段、胃肠道不同部位噬菌体群落对微生物群落的影响。

**5.4** 尽管病毒宏基因组学为研究人员提供了一种不需分离培养即可了解环境中全部病毒基因组信息的方法, 但仍需要以瘤胃优势菌群为宿主菌, 对噬菌体进行分离培养, 这将有助于全面了解噬菌体的生物学特性以及它们在维持瘤胃动态平衡中的作用。

**5.5** 关于瘤胃病毒的研究在国内处于亟待开发的阶段, 瘤胃是一个复杂、多变的微生态系统, 蕴藏着丰富的基因资源, 有大量新病毒等待研究人员去挖掘。故需要大力发展全基因组扩增技术和测序技

术,不断完善动物胃肠道病毒资源库,丰富瘤胃噬菌体基因组数据库,揭示噬菌体塑造瘤胃细菌群体结构的机制。

**5.6** 我国于 2020 年开始对饲料全面禁抗,鉴于瘤胃噬菌体对瘤胃微生物的裂解作用及对瘤胃营养物质循环的影响,噬菌体制剂作为抗生素替代品将成为畜牧界的重大课题。

#### 参考文献 (References):

- [1] WOMMACK K E, COLWELL R R. Virioplankton: Viruses in aquatic ecosystems[J]. *Microbiol Mol Biol Rev*, 2000, 64(1):69-114.
- [2] FOCARDI A, OSTROWSKI M, GOOSSEN K, et al. Investigating the diversity of marine bacteriophage in contrasting water masses associated with the East Australian Current (EAC) system[J]. *Viruses*, 2020, 12(3):317.
- [3] SHKOPOROV A N, RYAN F J, DRAPER L A, et al. Reproducible protocols for metagenomic analysis of human faecal phageomes [J]. *Microbiome*, 2018, 6(1):68.
- [4] BRAILSFORD M D, HARTMAN P A. Characterization of *Streptococcus durans* bacteriophages [J]. *Can J Microbiol*, 1968, 14(4):397-402.
- [5] HOOGENRAAD N J, HIRD F J R, HOLMES I, et al. Bacteriophages in rumen contents of sheep[J]. *J Gen Virol*, 1967, 1(4):575-576.
- [6] GILBERT R A, TOWNSEND E M, CREW K S, et al. Rumen virus populations: Technological advances enhancing current understanding [J]. *Front Microbiol*, 2020, 11:450.
- [7] KLIEVE A V, BAUCHOP T. Morphological diversity of ruminal bacteriophages from sheep and cattle[J]. *Appl Environ Microbiol*, 1988, 54(6):1637-1641.
- [8] MARTINEZ-MARTINEZ J, SWAN B K, WILSON W H. Marine viruses, a genetic reservoir revealed by targeted viromics [J]. *ISME J*, 2014, 8(5):1079-1088.
- [9] PAEZ-ESPINO D, ELOE-FADROSH E A, PAVLOPOULOS G A, et al. Uncovering earth's virome [J]. *Nature*, 2016, 536(7617):425-430.
- [10] SUTTON T D S, HILL C. Gut bacteriophage: Current understanding and challenges[J]. *Front Endocrinol*, 2019, 10:784.
- [11] MANRIQUE P, BOLDUC B, WALK S T, et al. Healthy human gut phageome[J]. *Proc Natl Acad Sci U S A*, 2016, 113(37):10400-10405.
- [12] DENMAN S E, MORGAVI D P, MCSWEENEY C S. Review: The application of omics to rumen microbiota function [J]. *Animal*, 2018, 12(S2):S233-S245.
- [13] MORGAVI D P, KELLY W J, JANSSEN P H, et al. Rumen microbial (meta)genomics and its application to ruminant production[J]. *Animal*, 2013, 7(S1):184-201.
- [14] BERGEN W G. Quantitative determination of rumen ciliate protozoal biomass with real-time PCR[J]. *J Nutr*, 2004, 134(12):3223-3224.
- [15] JANSSEN P H, KIRS M. Structure of the archaeal community of the rumen [J]. *Appl Environ Microbiol*, 2008, 74(12):3619-3625.
- [16] DENG W D, XI D M, MAO H M, et al. The use of molecular techniques based on ribosomal RNA and DNA for rumen microbial ecosystem studies; a review [J]. *Mol Biol Rep*, 2008, 35(2):265-274.
- [17] PARMAR N R, JAKHESARA S J, MOHAPATRA A, et al. Rumen virome: an assessment of viral communities and their functions in the rumen of an Indian buffalo[J]. *Curr Sci*, 2016, 111(5):919-925.
- [18] NAMONYO S, WAGACHA M, MAINA S, et al. A metagenomic study of the rumen virome in domestic caprids[J]. *Arch Virol*, 2018, 163(12):3415-3419.
- [19] BARYLSKI J, ENAULT F, DUTILH B E, et al. Analysis of spounaviruses as a case study for the overdue reclassification of tailed phages [J]. *Syst Biol*, 2020, 69(1):110-123.
- [20] ADAMS J C, GAZAWAY J A Jr, BRAILSFORD M D, et al. Isolation of bacteriophages from the bovine rumen[J]. *Experientia*, 1966, 22(11):717-718.
- [21] MILLER M E B, YEOMAN C J, CHIA N, et al. Phage-bacteria relationships and CRISPR elements revealed by a metagenomic survey of the rumen microbiome[J]. *Environ Microbiol*, 2012, 14(1):207-227.
- [22] SWAIN R A, NOLAN J V, KLIEVE A V. Natural variability and diurnal fluctuations within the bacteriophage population of the rumen [J]. *Appl Environ Microbiol*, 1996, 62(3):994-997.
- [23] ANDERSON C L, SULLIVAN M B, FERNANDO S C. Dietary energy drives the dynamic response of bovine rumen viral communities [J]. *Microbiome*,

- 2017,5(1):155.
- [24] ROSS E M, PETROVSKI S, MOATE P J, et al. Metagenomics of rumen bacteriophage from thirteen lactating dairy cattle [J]. *BMC Microbiol*, 2013, 13(1): 242.
- [25] POPOV V L, TESH R B, WEAVER S C, et al. Electron microscopy in discovery of novel and emerging viruses from the collection of the world reference center for emerging viruses and arboviruses (WRCEVA)[J]. *Viruses*, 2019, 11(5):477.
- [26] SU Z Q, TONG P P, ZHANG L, et al. First isolation and molecular characterization of blaCTX-M-121-producing *Escherichia coli* O157: H7 from cattle in Xinjiang, China[J]. *Front Vet Sci*, 2021, 8:574801.
- [27] KLIEVE A V, HEGARTY R S. Opportunities for biological control of ruminal methanogenesis[J]. *Aust J Agr Res*, 1991, 50: 1315-1319.
- [28] TAMADA H, HARASAWA R, SHINJO T. Isolation of a bacteriophage in *Fusobacterium necrophorum*[J]. *Nihon Juigaku Zasshi*, 1985, 47(3):483-486.
- [29] GILBERT R A, KELLY W J, ALTERMANN E, et al. Toward understanding phage: host interactions in the rumen; complete genome sequences of lytic phages infecting rumen bacteria[J]. *Front Microbiol*, 2017, 8:2340.
- [30] FRIEDERSDORFF J C A, KINGSTON-SMITH A H, PACHEBAT J A, et al. The isolation and genome sequencing of five novel bacteriophages from the rumen active against *Butyrivibrio fibrisolvens* [J]. *Front Microbiol*, 2020, 11:1588.
- [31] KLIEVE A V, GREGG K, BAUCHOP T. Isolation and characterization of lytic phages from *Bacterioides ruminicola ss brevis*[J]. *Curr Microbiol*, 1991, 23(4): 183-187.
- [32] CUERVO A, CARRASCOSA J L. Observation of bacteriophage ultrastructure by cryo-electron microscopy[J]. *Methods Mol Biol*, 2018, 1693:43-55.
- [33] HANDELSMAN J. Metagenomics; application of genomics to uncultured microorganisms[J]. *Microbiol Mol Biol Rev*, 2004, 68(4):669-685.
- [34] 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.
- HE B, TU C C. The advances and applications of viral metagenomics [J]. *Acta Veterinaria et Zootechnica Sinica*, 2012, 43(12):1865-1870. (in Chinese)
- [35] EISENHOFER R, MINICH J J, MAROTZ C, et al. Contamination in low microbial biomass microbiome studies; issues and recommendations [J]. *Trends Microbiol*, 2019, 27(2):105-117.
- [36] HUNGATE R E. The rumen and its microbes[M]. New York:Academic Press,1966:13-14.
- [37] 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.
- CHEN J H. Progress in the study of macrogenomics of environmental microorganisms [J]. *Jilin Agricultural*, 2010(6):44-45,49 (in Chinese).
- [38] CONCEIÇÃO-NETO N, ZELLER M, LEFRÈRE H, et al. Modular approach to customise sample preparation procedures for viral metagenomics; a reproducible protocol for virome analysis [J]. *Sci Rep*, 2015, 5:16532.
- [39] DUAN Y, LLORENTE C, LANG S, et al. Bacteriophage targeting of gut bacterium attenuates alcoholic liver disease[J]. *Nature*, 2019, 575(7783): 505-511.
- [40] CARROLL-PORTILLO A, COFFMAN C N, VARGA M G, et al. Standard bacteriophage purification procedures cause loss in numbers and activity[J]. *Viruses*, 2021, 13(2):328.
- [41] CORDOVA A, DESERNO M, GELBART W M, et al. Osmotic shock and the strength of viral capsids [J]. *Biophys J*, 2003, 85(1):70-74.
- [42] D' HUMIÈRES C, TOUCHON M, DION S, et al. A simple, reproducible and cost-effective procedure to analyse gut phageome; from phage isolation to bioinformatic approach [J]. *Sci Rep*, 2019, 9(1): 11331.
- [43] QIAO Y, LIU W L, LU N, et al. Rapid droplet multiple displacement amplification based on the droplet regeneration strategy [J]. *Anal Chim Acta*, 2021, 1141:173-179.
- [44] RUSSELL D A. Sequencing, assembling, and finishing complete bacteriophage genomes [J]. *Methods Mol Biol*, 2018, 1681:109-125.
- [45] ARDUI S, AMEUR A, VERMEESCH J R, et al. Single molecule real-time (SMRT) sequencing comes of age; applications and utilities for medical diagnostics [J]. *Nucleic Acids Res*, 2018, 46(5):2159-2168.
- [46] CAO J B, ZHANG Y Q, DAI M, et al. Profiling of human gut virome with Oxford nanopore technology [J]. *Med Microecol*, 2020, 4:100012.
- [47] ROUX S, ENAULT F, HURWITZ B L, et al. VirSorter: Mining viral signal from microbial genomic data [J]. *PeerJ*, 2015, 3:e985.

- [48] REN J, AHLGREN N A, LU Y Y, et al. VirFinder: A novel  $k$ -mer based tool for identifying viral sequences from assembled metagenomic data[J]. *Microbiome*, 2017, 5(1): 69.
- [49] AMGARTEN D, BRAGA L P P, DA SILVA A M, et al. MARVEL, a tool for prediction of bacteriophage sequences in metagenomic bins[J]. *Front Genet*, 2018, 9: 304.
- [50] ZHENG T T, LI J, NI Y Q, et al. Mining, analyzing, and integrating viral signals from metagenomic data [J]. *Microbiome*, 2019, 7(1): 42.
- [51] KIEFT K, ZHOU Z C, ANANTHARAMAN K. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences [J]. *Microbiome*, 2020, 8(1): 90.
- [52] GUO J R, BOLDUC B, ZAYED A A, et al. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses [J]. *Microbiome*, 2021, 9(1): 37.
- [53] SHARMA V K, KUMAR N, PRAKASH T, et al. MetaBioME: a database to explore commercially useful enzymes in metagenomic datasets[J]. *Nucleic Acids Res*, 2010, 38; D468-D472.
- [54] KANEHISA M, FURUMICHI M, TANABE M, et al. KEGG: new perspectives on genomes, pathways, diseases and drugs [J]. *Nucleic Acids Res*, 2017, 45(D1): D353-D361.
- [55] GALPERIN M Y, MAKAROVA K S, WOLF Y I, et al. Expanded microbial genome coverage and improved protein family annotation in the COG database[J]. *Nucleic Acids Res*, 2015, 43(D1): D261-D269.
- [56] FINN R D, COGGILL P, EBERHARDT R Y, et al. The Pfam protein families database: towards a more sustainable future [J]. *Nucleic Acids Res*, 2016, 44(D1): D279-D285.
- [57] ALCOCK B P, RAPHENYA A R, LAU T T Y, et al. CARD 2020: Antibiotic resistome surveillance with the comprehensive antibiotic resistance database[J]. *Nucleic Acids Res*, 2020, 48(D1): D517-D525.
- [58] YUTIN N, KAPITONOV V V, KOONIN E V. A new family of hybrid virophages from an animal gut metagenome[J]. *Biol Direct*, 2015, 10(1): 19.
- [59] SOLDEN L M, NAAS A E, ROUX S, et al. Interspecies cross-feeding orchestrates carbon degradation in the rumen ecosystem [J]. *Nat Microbiol*, 2018, 3(11): 1274-1284.
- [60] VAN ECK N J, WALTMAN L. Software survey: VOSviewer, a computer program for bibliometric mapping[J]. *Scientometrics*, 2010, 84(2): 523-538.
- [61] MARTINEZ-HERNANDEZ F, FORNAS O, GOMEZ M L, et al. Single-virus genomics reveals hidden cosmopolitan and abundant viruses[J]. *Nat Commun*, 2017, 8(1): 15892.
- [62] BREITBART M, SALAMON P, ANDRESEN B, et al. Genomic analysis of uncultured marine viral communities[J]. *Proc Natl Acad Sci U S A*, 2002, 99(22): 14250-14255.
- [63] GREGORY A C, ZAYED A A, CONCEIÇÃO-NETO N, et al. Marine DNA viral macro- and microdiversity from pole to pole [J]. *Cell*, 2019, 177(5): 1109-1123. e14.
- [64] CLOKIE M R J, MILLARD A D, LETAROV A V, et al. Phages in nature[J]. *Bacteriophage*, 2011, 1(1): 31-45.
- [65] GRUNINGER R J, RIBEIRO G O, CAMERON A, et al. Invited review: Application of meta-omics to understand the dynamic nature of the rumen microbiome and how it responds to diet in ruminants [J]. *Animal*, 2019, 13(9): 1843-1854.
- [66] EMERSON J B, ROUX S, BRUM J R, et al. Host-linked soil viral ecology along a permafrost thaw gradient[J]. *Nat Microbiol*, 2018, 3(8): 870-880.
- [67] BRUM J R, IGNACIO-ESPINOZA J C, KIM E H, et al. Illuminating structural proteins in viral “dark matter” with metaproteomics[J]. *Proc Natl Acad Sci U S A*, 2016, 113(9): 2436-2441.
- [68] ALLEN L Z, MCCROW J P, ININBERGS K, et al. The Baltic Sea virome: diversity and transcriptional activity of DNA and RNA viruses [J]. *mSystems*, 2017, 2(1): e00125-16.
- [69] BRUM J R, HURWITZ B L, SCHOFIELD O, et al. Seasonal time bombs: dominant temperate viruses affect Southern Ocean microbial dynamics[J]. *ISME J*, 2016, 10(2): 437-449.
- [70] KNOWLES B, SILVEIRA C B, BAILEY B A, et al. Lytic to temperate switching of viral communities[J]. *Nature*, 2016, 531(7595): 466-470.
- [71] SHKOPOROV A N, HILL C. Bacteriophages of the human gut: the “known unknown” of the microbiome [J]. *Cell Host Microbe*, 2019, 25(2): 195-209.

(编辑 范子娟)