

彭司华, 江守文, 孙 丹, 等. 宏基因组学技术在水产动物病毒鉴定中的应用[J]. 江苏农业学报, 2019, 35( 1 ): 229-237.  
doi: 10.3969/j.issn.1000-4440.2019.01.033

## 宏基因组学技术在水产动物病毒鉴定中的应用

彭司华<sup>1,2,3,4</sup>, 江守文<sup>2,3,4</sup>, 孙 丹<sup>2,3,4</sup>, 吴智超<sup>2,3,4</sup>, 陈 洁<sup>2,4</sup>

(1. 上海海洋大学水产与生命学院, 上海 201306; 2. 水产种质资源发掘与利用教育部重点实验室/上海海洋大学, 上海 201306; 3. 农业部国家水生动物病原库/上海海洋大学, 上海 201306; 4. 科学技术部海洋生物科学国际联合研究中心/上海海洋大学, 上海 201306)

**摘要:** 近年来水产养殖生产中病毒性疾病的发生越来越严重, 所以对水产动物病毒的深入研究成为水产研究者的一个重要课题。随着宏基因组学技术的不断发展, 以及测序成本的快速下降, 基于宏基因组学技术的水产动物病毒鉴定技术逐渐成为水产动物病毒鉴定的主要技术之一。本文总结了基于宏基因组学的水产动物病毒鉴定技术的主要方法, 详细综述了鉴定技术的实现流程, 评述了几种基于宏基因组学的水产动物病毒鉴定技术的优缺点, 并对基于宏基因组学的水产动物病毒鉴定技术的未来发展方向进行了展望。

**关键词:** 水产动物病毒; 宏基因组学; 病毒鉴定; 二代测序

**中图分类号:** S941.41 **文献标识码:** A **文章编号:** 1000-4440(2019)01-0229-09

## Application of metagenomics technology in aquaculture animal virus identification

PENG Si-hua<sup>1,2,3,4</sup>, JIANG Shou-wen<sup>2,3,4</sup>, SUN Dan<sup>2,3,4</sup>, WU Zhi-chao<sup>2,3,4</sup>, CHEN Jie<sup>2,4</sup>

(1. College of Fisheries and Life Science, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources/Shanghai Ocean University, Ministry of Education of China, Shanghai 201306, China; 3. National Pathogen Collection Center for Aquatic Animals, Ministry of Agriculture of China, Shanghai 201306, China; 4. International Research Center for Marine Biosciences at Shanghai Ocean University, Ministry of Science and Technology of China, Shanghai 201306, China)

**Abstract:** In recent years, the occurrence of viral diseases in the production of aquaculture has become more and more serious. Therefore, in-depth research on aquatic virus has become an important issue for aquatic researchers. With the continuous development of metagenomics technology and the rapid decline in sequencing costs, aquatic animal virus identification technology based on metagenomics technology has gradually become one of the main techniques for identification of aquatic animal viruses. This article summarized the main methods for the identification of aquatic animal viruses based on metagenomics, reviews the identification process in detail, and discussed the advantages and disadvantages of

the identification techniques. The development of the identification technology of aquatic animal viruses has been forecasted.

**Key words:** aquatic animal virus; metagenomics; virus identification; next generation sequencing

收稿日期: 2018-03-21

基金项目: 上海市自然科学基金项目(15ZR1420800); 上海海洋大学博士科研启动项目(A2-0203-00-100313); 上海海洋大学青年基金项目(A2-0203-00-100237)

作者简介: 彭司华(1962-), 男, 湖南古丈人, 博士, 副研究员, 主要从事生物信息学、基因组学研究。(E-mail) shpeng@shou.edu.cn

中国水产养殖产业飞速发展, 最近 29 年产量

一直居世界第一。据《2016年全国渔业经济统计公报》显示<sup>[1]</sup>,2016年中国水产品总产量达到 $6.9 \times 10^7$  t,其中养殖水产品产量 $5.1 \times 10^7$  t,占水产品总产量的74.5%。据联合国粮食及农业组织《世界渔业和水产养殖2016》报道<sup>[2]</sup>,中国的水产养殖产量占全球水产养殖产量的60%以上。然而,近些年来世界各国都遭遇到一个制约水产品产业发展的严重问题,即水产养殖生产中病毒性疾病的发生越来越严重。针对这一挑战,世界各国病毒学家都对水产动物病毒的研究给予了高度重视,并获得了大量的研究成果<sup>[3]</sup>。目前已从养殖和野生水生动物中鉴定出包括虹彩病毒、呼肠孤病毒、弹状病毒、疱疹病毒和双RNA病毒等不同病毒<sup>[4-5]</sup>。尤其是近年来,随着二代和三代测序组测序技术的发展,已经获得了许多水产动物病毒的全基因组序列,这给水产动物病毒的深入研究打下了良好的基础<sup>[6-7]</sup>。

宏基因组学(Metagenomics)是通过直接从环境样品中提取全部微生物的DNA,构建宏基因组文库,利用基因组学的研究策略研究样品所包含的全部微生物的基因组,并可刻画其中包含的微生物的群落功能。随着宏基因组学技术的深入发展,给水产动物病毒的鉴定与发现带来了很多便利<sup>[8-10]</sup>。基于宏基因组学发现病毒,过去10年来发表了很多十分有见地的综述<sup>[11-14]</sup>,但随着宏基因组学技术的进步,需要介绍最近几年来新的技术进展。2018年Greninger发表了关于基于宏基因组学发现RNA病毒的很深刻的综述<sup>[15]</sup>,但对于水产病毒领域的学者,更希望看到关于针对水产动物病毒更详尽的综述。尽管在水产动物病毒鉴定方面,Munang'andu最近发表了一篇较为详尽的综述<sup>[16]</sup>,但文章着重点在病毒发现的结果描述方面,而在介绍病毒鉴定的实现流程方面,则较少涉及。为弥补以上综述文献的不足,我们以水产动物中的鱼类为重点,着重介绍基于宏基因组的水产动物病毒鉴定实现流程的最新进展,并展望未来技术的发展方向。

本文按测序平台推出的先后顺序介绍6种测序平台,包括二代和三代测序技术。简要介绍各种测序平台在病毒鉴定中的应用。并介绍了利用新一代测序技术(Next generation sequencing,NGS。含二代和三代测序技术)用于水产动物病毒的鉴定案例,

其中以养殖和野生的鱼类病毒鉴定为重点。详细总结基于新一代测序技术发现的鱼类病毒的种类,所属病毒家族,宿主来源等信息。对基于新一代测序技术和宏基因组学技术进行水产动物病毒鉴定的详细操作流程进行综述。

## 1 基于宏基因组学水产养殖病毒鉴定的产生背景

传统的病毒鉴定方法主要有病毒培养、电镜观察、血清学反应和PCR方法等,操作繁琐,费时费力,且鉴定的效率较低。传统的方法在鉴定新病毒方面主要缺陷有:(1)很多病毒缺乏特异的细胞系,使得至少60%以上的病毒无法获得;(2)用电镜来观察病毒其灵敏性较低,许多情况下根本无法观察到;(3)由于高效的特异性抗体很难获得,用血清学反应来鉴定病毒也不理想;(4)而基于PCR法的病毒鉴定,必须首先得到病毒的DNA序列,这对于未知序列的病毒,以及变异性大的病毒的鉴定事实上很难实现<sup>[17]</sup>。

而基于二代测序的病毒宏基因组学方法直接利用鱼类样本中的病毒DNA序列作为研究对象。能够快速,有效地发现各种未知鱼类病毒,而且能直观地显示各种鱼类样本中病毒的种类与丰度。这样就可以方便地了解病毒的分布情况,对某些病原和潜在致病病毒的变化进行实时监测。同时有望诊断由新的或未知病原体引起的鱼类疾病<sup>[6]</sup>。

## 2 二代与三代测序技术简介<sup>[18]</sup>

### 2.1 454 测序或焦磷酸测序技术

这是基于宏基因组学病毒发现的第一种高通量测序技术。焦磷酸测序技术的原理是:引物与模板DNA退火后,在ATP硫酸化酶,DNA聚合酶,双磷酸酶和荧光素酶4种酶的共同作用下,将引物上每一个dNTP的聚合与一次荧光信号的释放偶联起来,实现实时测定DNA序列。454测序被广泛用于鉴定来自人和动物的新病毒<sup>[19]</sup>,例如虫媒病毒(Orbiviruses)<sup>[20]</sup>,冠状病毒<sup>[21]</sup>, $\gamma$ 乳头瘤病毒(Gammapapillomavirus)<sup>[22]</sup>和蝙蝠肝炎病毒和Pegivirus<sup>[23]</sup>等。虽然这种技术比Sanger测序有更高的通量,但是由于其成本高,错误率较高,且在测序通量方面比不上Illumina等技术,因此,该技术已被其他NGS技

术所取代。

## 2.2 Ion torrent 测序技术

这是 Life Technologies/Thermo Fisher Scientific 开发的一种基于半导体测序技术的测序平台,这种技术最早在 2012 发布,这个测序平台在基本原理上与 454 焦磷酸测序平台基本相似。这些快速的测序技术很快被应用到病毒的临床检测中,例如乙型肝炎病毒<sup>[24]</sup>, HIV 病毒<sup>[25-26]</sup>, HCV<sup>[27]</sup>。采用 Ion torrent 测序技术,对许多病毒进行了病毒基因组测序,例如正交病毒<sup>[28]</sup>、猪繁殖与呼吸综合征病毒<sup>[29]</sup>、多瘤病毒<sup>[23]</sup>、托斯卡纳病毒<sup>[30]</sup>、蓝舌病病毒<sup>[31]</sup>、轮状病毒<sup>[32]</sup>和流感病毒<sup>[33]</sup>等。

## 2.3 ABI SOLiD 测序技术

SOLiD 使用了连接法测序的策略,该测序平台通过检测“双碱基编码原理”的颜色编码序列的原理,实现了 DNA 序列的检测。在得到原始数据后的数据分析流程中,把原始颜色序列和转换成颜色编码的参考序列进行对比,从而实现把 SOLiD 颜色序列准确定位到参考序列上,从而实现测序的目的。连接测序所采用的底物是 8 个碱基荧光探针的混合物,只根据序列的位置,所测 DNA 就可以被探针准确标记。Solid 的缺点是读长只有 50~75 bp,但测序的精确度较高,可达 99.99%<sup>[34-35]</sup>。由于读长太短,现在在病毒鉴定中已经不采用这种测序技术。

## 2.4 Illumina 测序技术

这是 Solexa/Illumina 开发出的一个高通量测序平台, Illumina 测序技术采用边合成边测序的策略来实现测序。首先,在 DNA 片段两端加上通用接头来构建文库,然后把文库加载到 Flowcell 芯片上,经过桥式 PCR 反应后,对每一条文库片段都进行扩增,结果形成一个簇,通过边合成边测序反应,在经过多次循环后,最后获得完整的核酸序列<sup>[36]</sup>。采用这种测序技术发现的病毒有血小板减少症病毒<sup>[37]</sup>, 猴腺病毒<sup>[38]</sup>, 原病毒属病毒<sup>[39]</sup>等。

## 2.5 PacBio 测序技术

PacBio 测序技术是第三代基因测序技术,又被称为单分子实时 DNA 测序技术[Single molecule real time (SMRT<sup>TM</sup>) DNA sequencing]。该方法基于纳米孔的单分子读取技术,不需要扩增即可快速读取序列。Pacific Biosciences 的 SMRT 测序平台,被称为 RS II,该技术可产生长序列<sup>[40]</sup>。

PacBio 测序技术非常适合病毒宏基因组学研究,因为测序读长可以达到 10 kb 以上<sup>[41-43]</sup>。但是 PacBio 测序成本较高,且测序的错误率较高,测序精确度只能达到 86%左右。但是由于测序读长较长,这种技术将成为未来最主要的病毒鉴定测序技术之一。

## 2.6 Nanopore 测序技术

Nanopore 是一种单分子实时测序的新一代测序方法,其以单分子 DNA(RNA)通过生物纳米孔的电流变化推测碱基组成来进行测序。Nanopore 测序准确率和 PacBio 持平,为 86%左右。目前 Nanopore 的长序列主要应用于高质量基因组的组装,尤其是对于高杂合,高重复,大基因组等复杂基因组。2016 年前,主要用于一些小基因组如大肠杆菌,酵母,线虫等,目前已经应用在人、野生番茄、香蕉、欧洲鳗等多个复杂基因组组装上<sup>[44-46]</sup>。由于测序读长较长,这种技术也会成为未来病毒鉴定最重要的测序技术之一。

# 3 新一代测序技术和宏基因组学在水产新病毒发现上的应用

病毒宏基因组学已经成功用于新的水生动物病毒发现研究,包括鱼类、甲壳类动物、软体动物、海龟和海洋哺乳动物,本文以鱼类新病毒发现研究成果为重点进行评述。

2011 年以来,应用二代测序和宏基因组学技术,鉴定了许多养殖和野生鱼类病毒,例如虹彩病毒科病毒<sup>[20]</sup>。在大西洋三文鱼<sup>[47]</sup>、淡水鲤鱼<sup>[48]</sup>和金体美洲鲈鱼<sup>[49]</sup>体内鉴定了呼肠孤病毒,在欧洲鲈鱼体内鉴定了圆环病毒<sup>[50]</sup>,在鲤鱼<sup>[51]</sup>和金体美洲鲈鱼<sup>[52]</sup>体内发现了小 RNA 病毒科病毒,在鲤鱼体内鉴定了微小核糖核酸病毒属病毒<sup>[53]</sup>,在白亚口鱼和大太阳鱼体内鉴定了嗜肝 DNA 病毒<sup>[54-55]</sup>,在金体美洲鲈鱼体内鉴定了整体病毒科病毒<sup>[56]</sup>,在钓饵鱼体内鉴定了黑头豆科杯状病毒<sup>[57]</sup>。这些基于二代测序和宏基因组学鉴定的鱼类病毒案例的详细信息见表 1。

此外,二代测序技术和宏基因组学在甲壳动物虾<sup>[58-59]</sup>、贝类<sup>[60-61]</sup>、海星<sup>[62-63]</sup>、佛罗里达绿海龟<sup>[64]</sup>以及加利福尼亚海狮<sup>[65]</sup>、太平洋海豹<sup>[66]</sup>、白喙海豚<sup>[67]</sup>、海狗<sup>[68]</sup>等水产动物为宿主的病毒发现中也得到广泛应用。

表 1 采用二代测序技术(NGS)和宏基因组学分析发现的部分鱼类病毒

Table 1 Some fish viruses identification using next-generation sequencing (NGS) and metagenomic analysis

病毒英文名称	简称	病毒家族	类别	宿主	发现年份	文献
Erythrocytic necrosis virus	ENV	虹彩病毒科	dsDNA	太平洋鲱鱼	2011	[20]
Piscine orthoreovirus	PRV	呼肠孤病毒	dsRNA	大西洋三文鱼	2011	[47]
Catfish circovirus	CfCV	圆环病毒	ssDNA	欧洲鲶鱼	2013	[50]
Banna virus	BALV	呼肠孤病毒	dsRNA	淡水鲤鱼	2013	[48]
Carp picornavirus 1	CPV-1	小 RNA 病毒科	+ssRNA	鲤鱼	2014	[51]
Fathead minnow picornavirus	FHMV	小 RNA 病毒科	+ssRNA	金体美洲鲢鱼	2014	[52]
Fisavirus 1	FisaV1	微小核糖核酸病毒属	+ssRNA	鲤鱼	2015	[53]
White sucker hepatitis B virus	WSHBV	嗜肝 DNA 病毒	dsDNA	白亚口鱼	2015	[55]
Golden Shiner totivirus	GSTV	整体病毒科	dsRNA	金体美洲鲢鱼	2016	[56]
Piscine myocarditis like virus	PMCLV	呼肠孤病毒科	dsRNA	金体美洲鲢鱼	2016	[56]
Bluegill hepadnavirus	BGHB	嗜肝 DNA 病毒	dsDNA	大太阳鱼	2016	[54]
calicivirus	FHMCV-2012	黑头豆科杯状病毒	ssRNA	钓饵鱼	2017	[57]

#### 4 病毒鉴定流程

仅仅采用二代测序技术鉴定病毒是一个费时费力的工作,且要耗费较多的资金,现在已经较少采用这种方法来鉴定病毒了。随着宏基因组学技术的成熟,二代测序加上宏基因组学技术成为鉴定病毒的主流技术。由于绝大多数微生物无法人工培养,造成仅基于二代测序的方法鉴定病毒有很大的局限性。而基于宏基因组学技术可以克服大多数微生物不能培养的缺点,使得基于宏基因组学的病毒鉴定得到广泛采用。

一些研究者开发了一些基于宏基因组学的病毒鉴定流程软件(Pipeline for virus identification and discovery),例如中国疾控中心 Li 等开发的 VIP 病毒鉴定集成软件系统<sup>[69]</sup>,如图 1 所示。

Lin 等则开发了另一种集成软件系统 drVM<sup>[70]</sup>。但这些集成软件系统虽然使得系统容易使用,但它们往往只集成固定的几种软件。而实际上宏基因组学分析软件更新换代非常快,这就使得现有的集成软件系统的性能不一定是最好的。因此选择当时最好的软件来构建自己的分析流程是最好的选择。当然这需要较多的生物信息学软件使用经验。Tengs 等在大西洋三文鱼的病毒鉴定中提出了一个鱼类病毒的鉴定流程<sup>[71]</sup>,根据文中的描述,我们绘制出该病毒鉴定流程示意图(图 2)。在基本分析框架不变

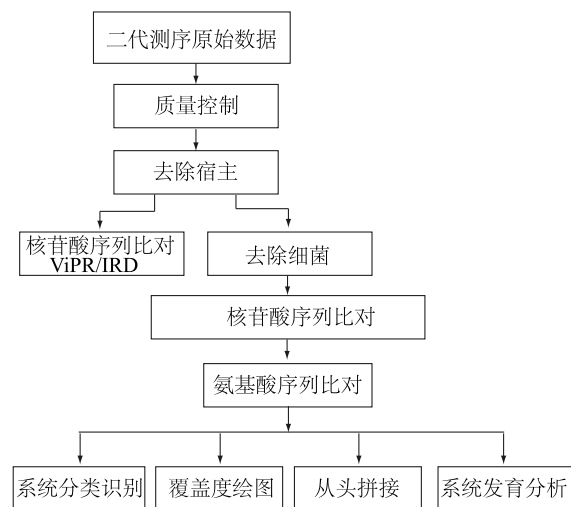


图 1 VIP 病毒鉴定集成软件系统示意图

Fig.1 Work flow of the VIP virus identification software system

的情况下,我们可以更换分析流程中的任何软件。例如该文作者采用了 Velvet<sup>[72]</sup> 软件来做宏基因组拼接,实际上我们可以把 Velvet 用其他宏基因组拼接软件来代替,例如 MEGAHIT<sup>[73]</sup>、Spades<sup>[14]</sup>、SOAPdenovo2<sup>[74]</sup> 等。质量控制工具 PRINSEQ<sup>[75]</sup> 也可以用 TRIMOMATIC<sup>[76]</sup>、NGS QC Toolkit<sup>[77]</sup> 等软件来代替。在对数据进行微生物分类学鉴定时采用的 METAXA2<sup>[78]</sup> 软件也可以换成普遍采用的同类型软件 QIIME<sup>[79]</sup>。



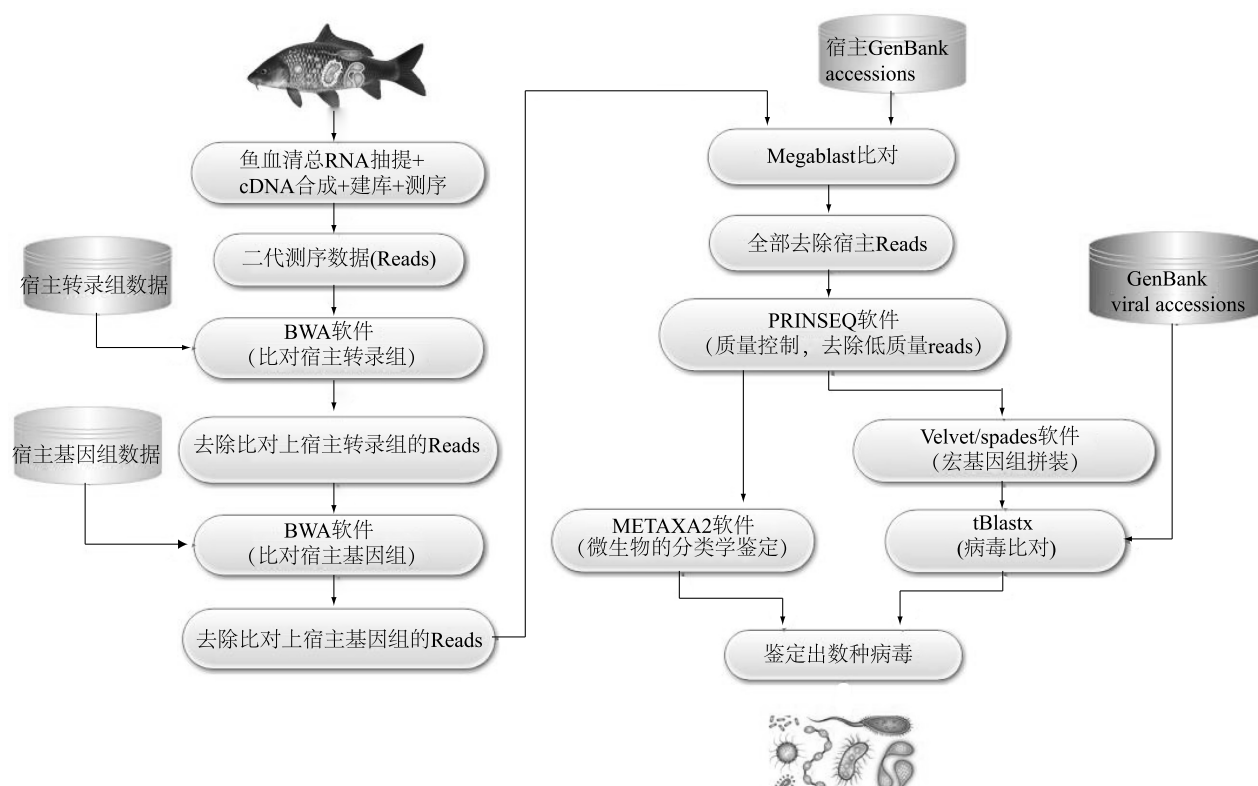


图2 鱼类病毒鉴定流程示意图

Fig.2 Schematic diagram of fish virus identification

我们课题组在研究青藏高原土壤与水体宏基因组研究中,采用 Torstein Tengs 等使用的 Pipeline 流程,成功地鉴定了一种病毒。此外,我们还开始了鲤鱼肠道宏基因组学研究,目前做的项目有鲤鱼 16S RNA 测序,鲤鱼肠道宏基因组测序。虽然目前我们测序的鲤鱼样本是健康的鲤鱼,但我们仍然开始了鲤鱼肠道病毒筛选工作。后续还打算对发病的鲤鱼样本进行深度测序,并初步打算增加三代测序方法,这样既可以拼接到更多的鲤鱼肠道微生物全长基因组,又可能筛选到鲤鱼肠道中的病毒。

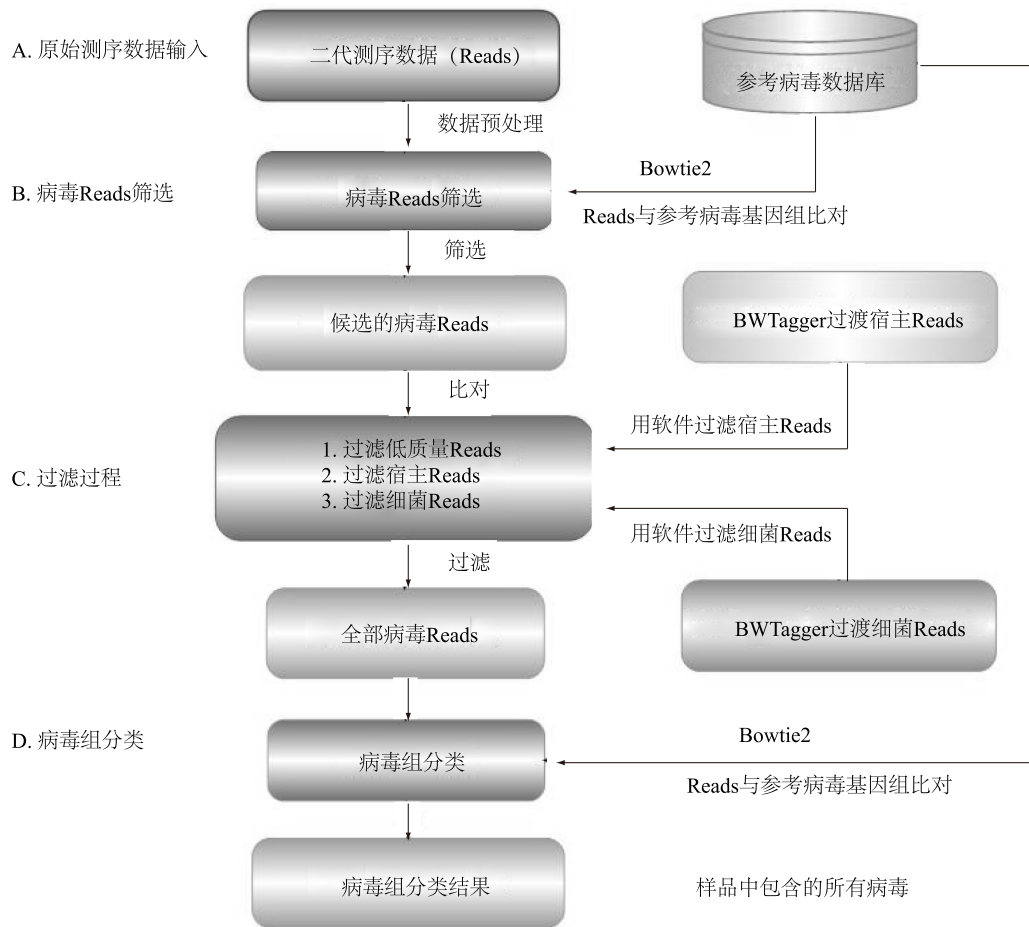
Rampelli 等开发的 ViromeScan<sup>[80]</sup>,其实现流程如 3 所示。这个软件只要输入二代测序 fastq 格式的原始数据 (Raw data),不需要拼接宏基因组,经过多步运算后就可以得到病毒组分类结果。其中采用的病毒组数据库可以在 NCBI 病毒基因组数据库<sup>[81]</sup>下载。在过滤宿主基因组和细菌 Reads 时,均采用了 BMTagger<sup>[82]</sup>。

最新报道的几种 Pipeline 病毒鉴定方法值得关

注, Ren 等推出了 VirFinder<sup>[83]</sup>。Lin 等开发了基于 web 的鉴定流程 Vipie<sup>[84]</sup>,另外一种新方法是 Tithi 等提出的基于单机版的病毒鉴定流程 FastVirome-Explorer<sup>[85]</sup>。这几种方法和上面介绍的 3 种方法实现流程虽有差别,但也基本相似,读者可以参考原文了解详细实现流程。

## 5 展望

本文综述了基于新一代测序和宏基因组学进行水产动物病毒鉴定和发现的基本原理和主要技术。尤其是随着近年来宏基因组学技术的完善以及测序成本的持续大幅度下降,使得宏基因组学技术已经成为水产动物疾病诊断和水产养殖新病原发现的一种重要工具。就新型病毒的发现而言,从装配方法 (De novo assemblies) 与生物信息学注释工具相结合,简化了鉴定与获得完整病毒基因组的步骤。因此,病毒宏基因组学可以用作新型病毒鉴定,病毒系统发育分析,病毒感染诊断,以及疾病监测的工具。



A: 单端或双端测序 fastq 格式数据; B: 过比对参考病毒组数据库序列来鉴定出候选病毒 reads; C: 过滤过程以去除低质量读数以及去除宿主、细菌与古细菌 reads; D: 留下的病毒 reads 再与病毒组数据库序列比对, 从而得到分类结果。

图 3 ViromeScan 软件的工作流程示意图

Fig.3 Workflow diagram of ViromeScan

在病毒宏基因组学分析中面临的一个重要挑战是, 使用宏基因组学技术得到的大部分病毒序列仍然未被注释, 这主要是因为它们与普通数据库中的任何已知序列没有相似性<sup>[86]</sup>。Mokili 等认为, 目前注释未知病毒序列不完善的主要原因还是缺乏合适的生物信息学分析工具<sup>[12]</sup>。Li 等指出, 病毒序列在装配中的问题之一是它们有时会形成嵌合重叠群, 这些嵌合重叠群在拼接时很难识别<sup>[69]</sup>。

病毒宏基因组学分析面临的另一个重要挑战是, 装配的病毒基因组的准确性还有待提高<sup>[12]</sup>。使用宏基因组学分析鉴定新型病毒的注释方法应力求避免出现假阳性<sup>[87]</sup>。鉴于大多数感染水生动物的病毒都没有被鉴定, 而且在数据库中也没有发现这种病毒, 所以新感染水生动物病毒的这种不准确性

可能更高。尽管如此, 我们认为基于基因组学的水产动物病毒鉴定将随着宏基因组学技术的不断成熟, 测序成本的不断下降, 以及已知病毒数据库的不断完善而得到更广泛的采用。

新型纳米孔测序技术 Nanopore 测序是一种非常有前途的测序技术。由于该测序仪的小型化以及样品准备极其简单, 且不需要借助碱基、多聚酶、连接酶等就能进行测序, 尤其适合野外携带。所以, Nanopore 测序技术有望在水产动物病毒鉴定方面得到更广泛的采用。

我们在表 1 中可以发现, 一种病毒可能感染多种鱼类, 例如呼肠孤病毒可以感染大西洋三文鱼、淡水鲤鱼、金体美洲鳊鱼等, 但感染的分子机制仍未揭示。随着宏基因组学的发展和完善, 更多的病毒序

列全长将会得到。这将给揭示病毒感染的机制带来很多方便,并将有力促进鱼病药物的研发,为更好地控制鱼病疫情打下基础。

### 参考文献:

- [1] 中国渔业报. 2016 年全国渔业经济统计公报 [EB/OL]. (2017-04-17) [2018-02-26]. <http://www.c-ampe.com/news/show.action?id=5d8c4fa896844e7cbbb0138aabcb2d43>.
- [2] 联合国粮食及农业组织. 世界渔业和水产养殖 [EB/OL]. (2016-09-27) [2018-02-25]. <http://www.fao.org/3/a-i5555c.pdf>.
- [3] 张奇亚, 桂建芳. 水生病毒学 [M]. 北京: 中国高等教育出版社, 2008: 3-6.
- [4] ZHANG Q, GUI J F. Virus genomes and virus-host interactions in aquaculture animals [J]. *Sci China Life Sci*, 2015, 58(2):156-169.
- [5] ZHANG Q Y, LI Z Q, GUI J F. Isolation of a lethal rhabdovirus from the cultured Chinese sucker *Myxocyprinus asiaticus* [J]. *Diseases of Aquatic Organisms*, 2000, 42(1):1-9.
- [6] ALAVANDI S V, POORNIMA M. Viral Metagenomics: a tool for virus discovery and diversity in aquaculture [J]. *Indian Journal of Virology*, 2012, 23(2):88-98.
- [7] CRANE M, HYATT A. Viruses of Fish: an overview of significant pathogens [J]. *Viruses-Basel*, 2011, 3(11):2025-2046.
- [8] EDWARDS R A, ROHWER F. Viral metagenomics [J]. *Nat Rev Microbiol*, 2005, 3(6):504-510.
- [9] HANDELSMAN J. Metagenomics: application of genomics to uncultured microorganisms [J]. *Microbiol Mol Biol Rev*, 2004, 68(4):669-685.
- [10] RIESENFELD C S, SCHLOSS P D, HANDELSMAN J. Metagenomics: genomic analysis of microbial communities [J]. *Annual Review of Genetics*, 2004, 38:525-552.
- [11] BEXFIELD N, KELLAM P. Metagenomics and the molecular identification of novel viruses [J]. *Veterinary Journal*, 2011, 190(2):191-198.
- [12] MOKILI J L, ROHWER F, DUTILH B E. Metagenomics and future perspectives in virus discovery [J]. *Current Opinion in Virology*, 2012, 2(1):63-77.
- [13] ROSARIO K, BREITBART M. Exploring the viral world through metagenomics [J]. *Current Opinion in Virology*, 2011, 1(4):289-297.
- [14] TANG P, CHIU C. Metagenomics for the discovery of novel human viruses [J]. *Future Microbiology*, 2010, 5(2):177-189.
- [15] GRENINGER A L. A decade of RNA virus metagenomics is (not) enough [J]. *Virus Research*, 2018, 244:218-229.
- [16] MUNANG'ANDU H M, MUGIMBA K K, BYARUGABA D K, et al. Current advances on virus discovery and diagnostic role of viral metagenomics in aquatic organisms [J]. *Frontiers in Microbiology*, 2017, 8:406.
- [17] 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.
- [18] SHENDURE J, JI H. Next-generation DNA sequencing [J]. *Nature Biotechnology*, 2008, 26(10):1135-1145.
- [19] DAY J M, BALLARD L L, DUKE M V, et al. Metagenomic analysis of the turkey gut RNA virus community [J]. *Virology Journal*, 2010, 7:313.
- [20] EMMENEGGER E J, GLENN J A, WINTON J R, et al. Molecular identification of erythrocytic necrosis virus (ENV) from the blood of Pacific herring (*Clupea pallasii*) [J]. *Veterinary Microbiology*, 2014, 174(1/2):16-26.
- [21] HONKAVUORI K S, BRIESE T, KRAUSS S, et al. Novel coronavirus and astrovirus in Delaware bay shorebirds [J]. *PLoS ONE*, 2014, 9(4):e93395.
- [22] PHAN T G, VO N P, ARONEN M, et al. Novel human gammapapillomavirus species in a nasal swab [J]. *Genome Announc*, 2013, 1(2):e0002213.
- [23] QUAN P-L, FIRTH C, CONTE J M, et al. Bats are a major natural reservoir for hepaciviruses and pegiviruses [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2013, 110(20):8194-8199.
- [24] YAN L, ZHANG H, MA H, et al. Deep sequencing of hepatitis B virus basal core promoter and precore mutants in HBeAg-positive chronic hepatitis B patients [J]. *Scientific Reports*, 2015, 5:1-9.
- [25] ARCHER J, WEBER J, HENRY K, et al. Use of four next-generation sequencing platforms to determine HIV-1 coreceptor tropism [J]. *PLoS ONE*, 2012, 7(11):e49602.
- [26] GIBSON R M, MEYER A M, WINNER D, et al. Sensitive deep-sequencing-based HIV-1 genotyping assay to simultaneously determine susceptibility to protease, reverse transcriptase, integrase, and maturation inhibitors, as well as HIV-1 coreceptor tropism [J]. *Antimicrobial Agents and Chemotherapy*, 2014, 58(4):2167-2185.
- [27] GASPARETO K V, RIBEIRO R M, MALTA F D M, et al. Resistance-associated variants in HCV subtypes 1a and 1b detected by Ion Torrent sequencing platform [J]. *Antiviral Therapy*, 2016, 21(8):653-660.
- [28] STEYER A, GUTIERREZ-AGUIRE I, KOLENC M, et al. High similarity of novel orthoreovirus detected in a child hospitalized with acute gastroenteritis to mammalian orthoreoviruses found in bats in Europe [J]. *Journal of Clinical Microbiology*, 2013, 51(11):3818-3825.
- [29] KVISGAARD L K, HJULSAGER C K, FAHNOE U, et al. A fast and robust method for full genome sequencing of porcine reproductive and respiratory syndrome virus (PRRSV) Type 1 and Type 2 [J]. *Journal of Virological Methods*, 2013, 193(2):697-705.
- [30] NOUGAIREDE A, BICHAUD L, THIBERVILLE S-D, et al. Isolation of toscana virus from the cerebrospinal fluid of a man with meningitis in Marseille, France, 2010 [J]. *Vector-Borne and Zoonotic Diseases*, 2013, 13(9):685-688.

- [31] LORUSSO A, MARCACCI M, ANCORA M, et al. Complete genome sequence of bluetongue virus serotype 1 circulating in Italy, obtained through a fast next-generation sequencing protocol [J]. *Genome Announc*, 2014, 2(1):e00093-14.
- [32] NYAGA M M, STUCKER K M, ESONA M D, et al. Whole-genome analyses of DS-1-like human G2P 4 and G8P 4 rotavirus strains from Eastern, Western and Southern Africa [J]. *Virus Genes*, 2014, 49(2):196-207.
- [33] VAN DEN HOECKE S, VERHELST J, VUYLSTEKE M, et al. Analysis of the genetic diversity of influenza A viruses using next-generation DNA sequencing [J]. *Bmc Genomics*, 2015, 16:79.
- [34] KIRCHER M, KELSO J. High-throughput DNA sequencing - concepts and limitations [J]. *Bioessays*, 2010, 32(6):524-536.
- [35] MOROZOVA O, MARRA M A. Applications of next-generation sequencing technologies in functional genomics [J]. *Genomics*, 2008, 92(5):255-264.
- [36] ZHANG P, FERNANDES H. Comparison of two next-generation sequencing systems: assessment of efficiency and quality of results [J]. *Journal of Molecular Diagnostics*, 2015, 17(6):844.
- [37] XU B, LIU L, HUANG X, et al. Metagenomic analysis of fever, thrombocytopenia and leukopenia syndrome (fls) in henan province, China: discovery of a new bunyavirus [J]. *PLoS Pathogens*, 2011, 7(11):e1002369.
- [38] CHIU C Y, YAGI S, LU X Y, et al. A novel adenovirus species associated with an acute respiratory outbreak in a baboon colony and evidence of coincident human infection [J]. *mBio*, 2013, 4(2):e00084.
- [39] PHAN T G, DRENO B, DA COSTA A C, et al. A new protoparvovirus in human fecal samples and cutaneous T cell lymphomas (mycosis fungoides) [J]. *Virology*, 2016, 496:299-305.
- [40] REUTER J A, SPACEK D V, SNYDER M P. High-throughput sequencing technologies [J]. *Mol Cell*, 2015, 58(4):586-597.
- [41] BULL R A, ELTAHLA A A, RODRIGO C, et al. A method for near full-length amplification and sequencing for six hepatitis C virus genotypes [J]. *Bmc Genomics*, 2016, 17:247.
- [42] HO C K Y, RAGHWANI J, KOEKKOEK S, et al. Characterization of hepatitis C virus (HCV) envelope diversification from acute to chronic infection within a sexually transmitted HCV cluster by using single-molecule, real-time sequencing [J]. *Journal of Virology*, 2017, 91(6):e02262.
- [43] LI J, WANG M, YU D, et al. A comparative study on the characterization of hepatitis B virus quasispecies by clone-based sequencing and third-generation sequencing [J]. *Emerging Microbes & Infections*, 2017, 6:e100.
- [44] MAGI A, SEMERARO R, MINGRINO A, et al. Nanopore sequencing data analysis: state of the art, applications and challenges [J]. *Brief Bioinform*, 2017, doi: 10.1093/bib/bbx1062.
- [45] SCHMIDT M H W, VOGEL A, DENTON A K, et al. De novo assembly of a new solanum pennellii accession using Nanopore sequencing [J]. *Plant Cell*, 2017, 29(10):2336-2348.
- [46] JAIN M, KOREN S, MIGA K H, et al. Nanopore sequencing and assembly of a human genome with ultra-long reads [J]. *Nat Biotechnol*, 2018, 36(4):338-345.
- [47] PALACIOS G, LOVOLL M, TENGS T, et al. Heart and skeletal muscle inflammation of farmed salmon is associated with infection with a novel reovirus [J]. *PLoS ONE*, 2010, 5(7):e11487.
- [48] REUTER G, BOROS A, DELWART E, et al. Novel seadornavirus (family reoviridae) related to Banna virus in Europe [J]. *Archives of Virology*, 2013, 158(10):2163-2167.
- [49] MOR S K, PHELPS N B D. Detection and molecular characterization of a novel piscine-myocarditis-like virus from baitfish in the USA [J]. *Archives of Virology*, 2016, 161(7):1925-1931.
- [50] LORINCZ M, DAN A, LANG M, et al. Novel circovirus in European catfish (silurus glanis) [J]. *Archives of Virology*, 2012, 157(6):1173-1176.
- [51] LANGE J, GROTH M, FICHTNER D, et al. Virus isolate from carp: genetic characterization reveals a novel picornavirus with two aphthovirus 2A-like sequences [J]. *Journal of General Virology*, 2014, 95:80-90.
- [52] PHELPS N B D, MOR S K, ARMIEN A G, et al. Isolation and molecular characterization of a novel picornavirus from baitfish in the USA [J]. *PLoS ONE*, 2014, 9(2):e87593.
- [53] REUTER G, PANKOVICS P, DELWART E, et al. A novel posavirus-related single-stranded RNA virus from fish (cyprinus carpio) [J]. *Archives of Virology*, 2015, 160(2):565-568.
- [54] DILL J A, CAMUS A C, LEARY J H, et al. Distinct viral lineages from fish and amphibians reveal the complex evolutionary history of Hepadnaviruses [J]. *Journal of Virology*, 2016, 90(17):7920-7933.
- [55] HAHN C M, IWANOWICZ L R, CORNMANN R S, et al. Characterization of a novel Hepadnavirus in the white Sucker (catostomus commersonii) from the Great Lakes region of the United States [J]. *J Virol*, 2015, 89(23):11801-11811.
- [56] MOR S K, PHELPS N B D. Molecular detection of a novel totivirus from golden shiner (notemigonus crysoleucas) baitfish in the USA [J]. *Archives of Virology*, 2016, 161(8):2227-2234.
- [57] THOMPSON L R, SANDERS J G, MCDONALD D, et al. A communal catalogue reveals Earth's multiscale microbial diversity [J]. *Nature*, 2017, 551(7681):457-463.
- [58] PHAM H T, YU Q, BOISVERT M, et al. A circo-like virus isolated from Penaeus monodon shrimps [J]. *Genome Announc*, 2014, 2(1):e01172.
- [59] ZHOU J, ZHANG W, YAN S, et al. Diversity of virophages in metagenomic data sets [J]. *Journal of Virology*, 2013, 87(8):4225-4236.
- [60] DAYARAM A, GOLDSTIEN S, ZAWAR-REZA P, et al. Novel ssDNA virus recovered from estuarine Mollusc (amphibola crenata) whose replication associated protein (Rep) shares similarities with Rep-like sequences of bacterial origin [J]. *Journal of General Virology*, 2013, 94:1104-1110.



- [61] DAYARAM A, GOLDSTIEN S, ZAWAR-REZA P, et al. Identification of Starling Circovirus in an estuarine mollusc (*amphibola crenata*) in New Zealand using metagenomic approaches [J]. *Genome Announc*, 2013, 1(3):e00278.
- [62] FAHSBENDER E, HEWSON I, ROSARIO K, et al. Discovery of a novel circular DNA virus in the Forbes sea star, *Asterias forbesi* [J]. *Archives of Virology*, 2015, 160(9):2349-2351.
- [63] HEWSON I, BUTTON J B, GUDENKAUF B M, et al. Densovirus associated with sea-star wasting disease and mass mortality [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2014, 111(48):17278-17283.
- [64] NG T F F, MANIRE C, BORROWMAN K, et al. Discovery of a novel single-stranded DNA virus from a sea Turtle fibropapilloma by using viral metagenomics [J]. *Journal of Virology*, 2009, 83(6):2500-2509.
- [65] NG T F F, SUEDEMEYER W K, WHEELER E, et al. Novel anellovirus discovered from a mortality event of captive California sea lions [J]. *Journal of General Virology*, 2009, 90:1256-1261.
- [66] NG T F F, WHEELER E, GREIG D, et al. Metagenomic identification of a novel anellovirus in Pacific harbor seal (*Phoca vitulina richardsii*) lung samples and its detection in samples from multiple years [J]. *Journal of General Virology*, 2011, 92:1318-1323.
- [67] SIEGERS J Y, VAN DE BILDT M W G, VAN ELK C E, et al. Genetic relatedness of dolphin rhabdovirus with fish rhabdoviruses [J]. *Emerging Infectious Diseases*, 2014, 20(6):1081-1082.
- [68] KLUGE M, CAMPOS F S, TAVARES M, et al. Metagenomic survey of viral diversity obtained from feces of subantarctic and south American fur Seals [J]. *PLoS ONE*, 2016, 11(3):e0151921.
- [69] LI Y, WANG H, NIE K, et al. VIP: an integrated pipeline for metagenomics of virus identification and discovery [J]. *Scientific Reports*, 2016, 6:23774.
- [70] LIN H-H, LIAO Y C. drVM: a new tool for efficient genome assembly of known eukaryotic viruses from metagenomes [J]. *Gigascience*, 2017, doi:10.1093/gigascience/gix1003.
- [71] TENGS T, RIMSTAD E. Emerging pathogens in the fish farming industry and sequencing based pathogen discovery [J]. *Developmental and Comparative Immunology*, 2017, 75:109-119.
- [72] ZERBINO D R, BIRNEY E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs [J]. *Genome Research*, 2008, 18(5):821-829.
- [73] LI D, LIU C M, LUO R, et al. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph [J]. *Bioinformatics*, 2015, 31(10):1674-1676.
- [74] LUO R B, LIU B H, XIE Y L, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler [J]. *Gigascience*, 2012, doi:10.1186/2047-1217x-1181-1118.
- [75] SCHMIEDER R, EDWARDS R. Quality control and preprocessing of metagenomic datasets [J]. *Bioinformatics*, 2011, 27(6):863-864.
- [76] BOLGER A M, LOHSE M, USADEL B. Trimmomatic: a flexible trimmer for Illumina sequence data [J]. *Bioinformatics*, 2014, 30(15):2114-2120.
- [77] PATEL R K, JAIN M. NGS QC Toolkit: A Toolkit for quality control of next generation sequencing data [J]. *PLoS ONE*, 2012, 7(2):e30619.
- [78] BENGTSOON-PALME J, HARTMANN M, ERIKSSON K M, et al. metaxa2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data [J]. *Molecular Ecology Resources*, 2015, 15(6):1403-1414.
- [79] CAPORASO J G, KUCZYNSKI J, STOMBAUGH J, et al. QIIME allows analysis of high-throughput community sequencing data [J]. *Nature Methods*, 2010, 7(5):335-336.
- [80] RAMPELLI S, SOVERINI M, TURRONI S, et al. ViromeScan: a new tool for metagenomic viral community profiling [J]. *Bmc Genomics*, 2016, doi:10.1186/s12864-016-12446-12863.
- [81] NCBI. TNvg [EB/OL]. (2015-02-12) [2018-02-25]. <http://www.ncbi.nlm.nih.gov/genomes/GenomesGroup.cgi?opt=virus&taxid=10239>.
- [82] ROTMISTROVSHY K. BMTagger [EB/OL]. (2012-09-02) [2018-02-25]. <ftp://ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger/>.
- [83] 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, doi:10.1186/s40168-0017-40283-40165.
- [84] LIN J, KRAMNA L, AUTIO R, et al. Vipie: web pipeline for parallel characterization of viral populations from multiple NGS samples [J]. *Bmc Genomics*, 2017, doi:10.1186/s12864-017-13721-12867.
- [85] TITHI S S, AYLWARD F O, JENSEN R V, et al. Fast virome explorer: a pipeline for virus and phage identification and abundance profiling in metagenomics data [J]. *PeerJ*, 2018, 6:e4227.
- [86] ROSARIO K, NILSSON C, LIM Y W, et al. Metagenomic analysis of viruses in reclaimed water [J]. *Environmental Microbiology*, 2009, 11(11):2806-2820.
- [87] BIBBY K. Metagenomic identification of viral pathogens [J]. *Trends Biotechnol*, 2013, 31(5):11-15.

(责任编辑:陈海霞)