

基于高通量测序技术的食源性病原体检测 应用进展

戴陈伟, 许 勇, 武昌俊, 蔡 标*

(安徽省医学科学研究院, 合肥 230061)

摘 要: 高通量测序是 DNA 测序技术发展中里程碑式的突破。面对突然爆发的食源性病原体疫情, 食源性病原体检测正从传统的物理化学方法以及核酸扩增杂交技术向无需富集分离的高通量测序技术发展。高通量测序技术把食源性病原体看成一个整体, 直接对食品样本中的所有病原体进行全基因组测序, 得到病原体基因组数据, 并进一步通过生物信息学分析, 得到病原体的基因型、毒力和耐药性报告。高通量测序技术广泛应用于食源性病原体检测的难点在于对测序数据的准确分析, 但其在食源性病原体的快速检测、预防食源疫情传播和日常的食源疫情监测, 尤其是发现未知食源性病原体方面拥有巨大的潜力。本文主要对基于高通量测序技术的食源性病原体检测技术的原理、应用进行综述, 并介绍了现阶段面临的挑战和机遇。

关键词: 食源性病原体; 高通量测序技术; 全基因组测序; 数据分析

Application progress in food-borne pathogens detection based on high-throughput sequencing technology

DAI Chen-Wei, XU Yong, WU Chang-Jun, CAI Biao *

(Anhui Academy of Medical Sciences, Hefei 230061, China)

ABSTRACT: High throughput sequencing technology is a milestone breakthrough in the development of DNA sequencing technology. In face of sudden outbreaks of foodborne pathogens, the detection of foodborne pathogens is developing from traditional methods such as physical and chemical methods, nucleic acid amplification and hybridization to high-throughput sequencing technology without enrichment and separation. High-throughput sequencing technology regards foodborne pathogens as a whole and directly sequences the whole genomes of all pathogens in food samples to obtain and analyze the genome data of pathogens, in order to get the genotype, virulence and drug resistance reports of pathogens. Difficulties of high-throughput sequencing technology applying in Food-borne Pathogens detection lie in the accurate analysis of sequencing data, but it has great potential in fast detection of food-borne pathogens, prevention and daily surveillance of food-borne epidemics, especially in the detection of unknown food-borne pathogens. This paper reviewed the principle and application of foodborne pathogen detection technology based on high-throughput sequencing technology, and introduced the challenges and opportunities facing the current stage.

KEY WORDS: foodborne pathogens; high-throughput sequencing technology; whole-genome sequencing; data analysis

基金项目: 安徽省十三五医疗卫生重点专科建设项目(皖卫科教[2017]30号)、2018年度安徽省卫计委科研计划项目(2018YK008)

Fund: Supported by "13th Five-Year Plan" Key Medical and Health Specialty of Anhui (Anhui Health Science and Education [2017] 30), and Health and Family Planning Commission of Anhui Province in 2018 Project (2018YK008)

***通讯作者:** 蔡标, 副研究员, 主要研究方向为病原微生物和流行病学。E-mail: caib81@163.com

***Corresponding author:** CAI Biao, Associate Professor, the Medical Science Institute of Anhui Province, No.15, Yonghong Road, Lu yang District, Hefei 230061, China. E-mail: caib81@163.com

1 引言

食品安全历来是关乎人民健康与社会和谐的热点问题, 食源性病原体(foodborne pathogens, FBPs)是一类可以引起食物中毒或以食品为传播媒介的致病性病原体, 包括食源性的细菌、病毒、真菌和寄生虫等^[1-3], 它们直接或经过传染媒介引起食源性疾病, 危害人类健康, 并造成极大的经济损失^[4]。正是由于引起食源性疾病的病原体种类繁多, 且部分病毒类病原体变异速度较快^[5], 如何对 FBPs 进行快速准确的诊断, 成为 FBPs 检测领域研究的重点。传统 FBPs 的检测技术, 如涂片镜检^[6]、生化检测^[7]、免疫试验^[8]等, 通常需要耗费时间和人力来对样本中的病原体进行分离和富集, PCR 检测^[9]和基因芯片检测^[10]需要对样本中可能存在的病原体基因组的先验知识有所了解, 因而难以对未知微生物或混合病原体进行准确鉴定。自 2005 年, 454 生命技术公司革命性地推出高通量测序系统以来^[11], 高通量测序(high-throughput sequencing, HTS)技术越来越多地在未知基因组信息的情况下应用于病原体的检测^[12,13]。近年来, 利用 HTS 技术开展的全基因组测序(whole-genome sequencing, WGS)已被广泛用于提供 FBPs 的详细特征, 为沙门氏菌、大肠杆菌、李斯特菌、诺如病毒等 FBPs 的遗传组成提供了深刻的见解^[14], 基于 HTS 技术的食源性病原体检测技术也得到了快速的发展^[15]。

本研究阐述了运用 HTS 技术检测食源性病原体的原理, 高通量实验数据的处理流程, 以及基于 HTS 技术在检测食源性细菌、病毒、真菌和寄生虫类的应用, 并展望了现阶段基于 HTS 技术的食源性病原体检测技术面临的机遇与挑战, 为高通量测序技术在检测食源性病原体上的进一步应用提供参考。

2 基于 HTS 技术的食源性病原体检测的原理

HTS 技术是对传统测序技术的一次革命性的改变^[16], 其在生物学和医学的各个领域都开始发挥着重要作用。Illumina 公司的 Solexa 测序是目前最流行的 HTS 技术, 其技术原理是采用边合成边测序反应, 在合成反应的碱基延伸过程中, 将 4 种不同的脱氧核糖核苷三磷酸(dNTP)标记上 4 种不同的荧光信号, 每个循环只按照碱基配对原则延伸一个碱基, 按照荧光信号确认每次延伸的碱基种类, 最终完整读取每个 DNA 片段的序列^[16]。Solexa 测序读取片段为相对较短(100~250 bp), 碱基读取错误率低。在 Solexa 测序流行之前, 罗氏 454 测序是主导技术, 其技术原理也是采用边合成边测序反应, 但只依次向反应体系中加入一种 dNTP, 如果某种 dNTP 刚好可以和文库 DNA 片段的下一个碱基配对, DNA 聚合酶将会把它添加到测序引物的 3' 末端, 并通过后续反应释放出光信号, 这样, 每当一个碱

基与模板进行配对, 就会释放出一分子的光信号, 一一对应就可以读取文库序列。454 测序读取片段较长(一般 ≤ 700 bp), 但碱基读取错误率相对较高。ABI SOLID 测序技术是另一个具有代表性的高通量测序技术, 基于该系统采用的双碱基编码技术, 在测序过程中会对每个碱基检测 2 次, 其检测准确性接近 99.95%, 因而能够区分测序错误和单核苷酸多态性, 主要适用于单核苷酸多态性的检测, 但该测序技术的片段长短约 100 bp, 后续的序列拼接工作相对复杂^[16]。任何平台的 HTS 技术未必适合每一个测序项目, 我们在开展实验时应根据病原体基因组研究现状和项目测序通量的要求等加以选择(表 1)^[16,17]。

一个食品样本无需经过病原体富集培养、无需对病原体基因组具有先验知识、同时也只需要少许的起始核酸量, 经过高通量测序可以产生数千到数百万个序列片段, 对应数亿个碱基的序列, 利用生物信息学软件和核酸数据库中的参考基因组信息对所得到的序列进行过滤筛选和比对分析^[17], 即可从单个样本海量的扩增序列中拼接获得高质量的基因组图谱, 从而进一步得到病原体的基因型、毒力和耐药性报告等。

3 基于 HTS 技术的食源性病原体检测的数据处理流程

根据样本性质、来源、提取方法以及建立文库策略的不同, 基于 HTS 技术的 FBPs 检测的数据处理方法也可能略有不同。FBPs 样本大多来自粮食、蔬菜、水产品、奶制品、水果、畜禽产品等, 在提取病原体的 DNA 或 RNA 后建立文库, 利用高通量测序仪测序, 可以获得存储格式为 fasta 或 fastq 的初始测序数据, 初始数据经过一系列生物信息学软件的分析处理得到样品中的病原体信息。我们将基于 HTS 技术的 FBPs 检测的数据处理流程(图 1, 附每步可用的分析软件)概括为以下 5 个步骤: (1) 测序数据的质量控制^[18,19], 目的是为了去除非目标序列、测序接头、引物序列以及过滤掉低质量的测序数据; (2) 与参考基因组比对, 以分拣测序片段^[20,21]; (3) 目的片段序列拼接^[22,23], 利用生物信息学软件拼接得到 configs 或 scaffolds; (4) configs/scaffolds 与参考序列比对^[24]; (5) 结果可视化^[25]与下游分析^[26,27], 获得全部病原体基因组信息, 完成病原体检测流程。

4 基于 HTS 技术的食源性病原体检测的应用

4.1 细菌类食源性病原体

细菌类 FBPs 是人类食物链中最常见的病原体, 主要包括大肠杆菌、沙门氏菌、志贺氏菌、葡萄球菌和李斯特菌等。细菌类 FBPs 具有传播快、在食品中含量少的特点, 因此, 需要我们有能力检测出在各种食品来源中极低水平存在的细菌类 FBPs。大部分细菌类 FBPs 已有较为深入的

表 1 3 种高通量测序技术的特点和技术优势
Table 1 Characteristics and technical advantages of 3 HTS technologies

测序技术	测序长度	测序保真度	技术优势
罗氏 454 测序技术	100~250 bp	较高	准确度高, 一致性好
Illumina Solexa 测序技术	≤700 bp	相对较低	需样品量少, 简单、快速
ABI SOLID 测序技术	约 100 bp	较高	通量高, 更适合单核苷酸多态性分析和甲基化分析

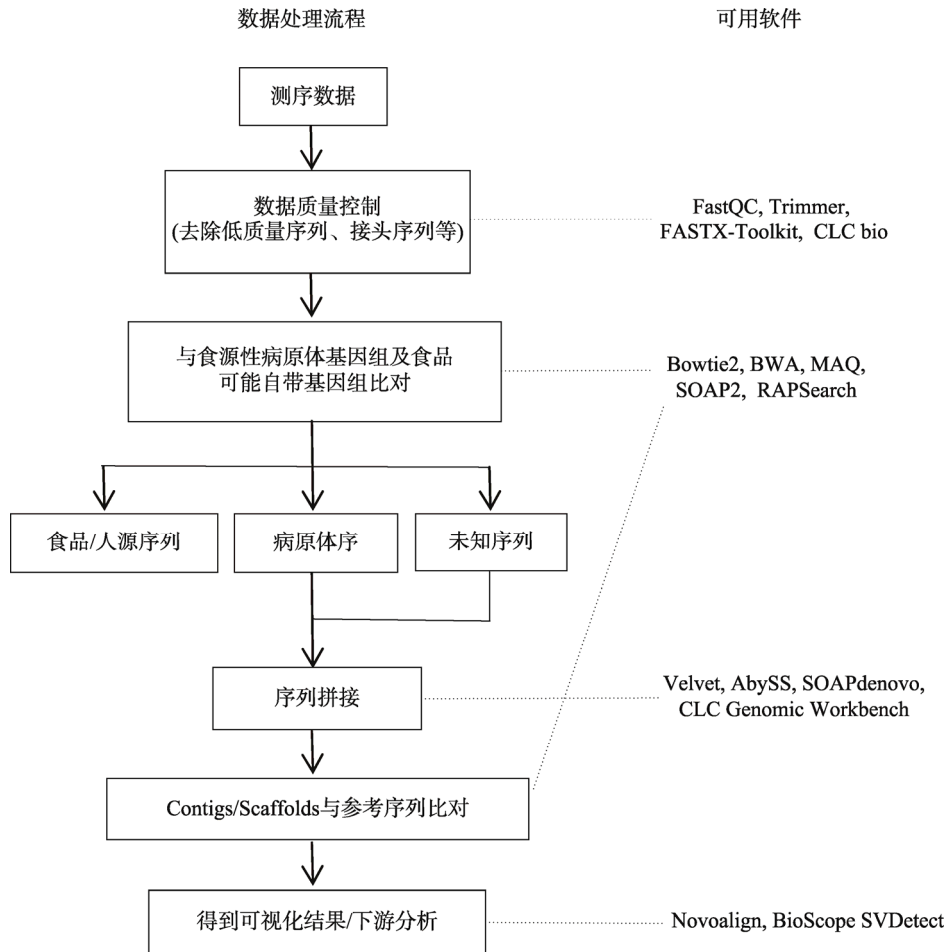


图 1 基于 HTS 技术的病原体检测的数据处理流程

Fig.1 Data processing flow of FBPs detection based on HTS technology

研究基础, 基因组信息也大多被上传到公共数据库, 因此, HTS 技术可以很好的应用到细菌类 FBPs 的鉴定、流行病学研究、进化分析和日常监测之中。

大肠杆菌 O157 是一种非常重要的 FBPs, 感染人后经常导致严重的出血性腹泻和腹部绞痛, 其致病菌主要为 O157:H7^[28]。Jenkins 等^[29]利用 HTS 技术成功鉴别出豆制品作为大肠杆菌 O157:H7 同时爆发的两处食品污染源, 采用全基因组测序与拼接, 辅以单核苷酸多态分析, 并与公共数据库中其他菌株的数据进行比较, 构建了系统进化发育树, 确认了感染爆发的菌株。彭昊等^[30]通过 HTS 技术对两株有毒力差异的大肠杆菌 O157:H7 进行鉴定, 采用测序

评估和基因功能注释等生物信息学方法对测序结果进行分析, 为下一步开展大肠杆菌 O157:H7 毒力相关基因的研究奠定了基础。单增李斯特菌感染可引起败血症、脑膜炎、肺炎、流产和死胎等症状, 是我国重点监测的 FBPs 之一^[31]。Allam 等^[32]运用 HTS 技术进行全基因组测序, 报告了发生在 2017 年至 2018 年期间, 与南非一次大规模李斯特菌爆发有关的 10 株单核细胞增多性李斯特菌 6 型分离株的全基因组序列。在我国, 沙门氏菌是细菌性食物中毒最常见的病原菌之一, 严重危害公众健康和生命^[33]。通常, 基于沙门氏菌富集培养的检测方法复杂而又耗时, Jarvis 等^[34]对 91 个茼蒿菜叶样品同时进行 16S rDNA 测序和全基

因组测序,发现利用细菌 16S rDNA 测序鉴定的方法对培养后沙门氏菌阳性的样品进行检测,只有较低沙门氏菌检出率,这可能是由于测序深度不够或是引物设计的不合理,而对沙门氏菌培养阳性的芜荑样品进行全基因组测序分析,沙门氏菌检出率 100%,且可以通过测序数据分析判断沙门氏菌污染程度的不同。

4.2 病毒类食源性病原体

病毒类 FBPs 主要包括甲型肝炎病毒、诺如病毒、轮状病毒、戊型肝炎病毒和腺病毒等,主要引起流行范围广、传染性强的胃肠道、肝脏以及神经系统疾病。病毒类 FBPs 常常难以直接确诊,因为病毒通常含量较低,大多难以培养,少数可以培养的病毒,除了需要非常合适的宿主,还需要其它严苛的条件^[35]。作为在全球流行范围最广的病毒类 FBPs,诺如病毒、甲型肝炎病毒、戊型肝炎病毒都是无包膜的单链 RNA 病毒,主要通过食物进行传播^[36]。严重急性呼吸综合征(Severe acute respiratory syndrome, SARS)病毒、中东呼吸综合征(Middle East Respiratory Syndrome, MERS)病毒和埃博拉病毒则是人畜共患的 RNA 病毒,也有可能以食物为媒介进行传播^[37]。与 DNA 病毒相比, RNA 病毒缺乏具有错误修复机制的聚合酶,因此 RNA 在复制过程中出现错误时无法及时修正,这就导致了 RNA 病毒较快的进化速度,新型 RNA 病毒的爆发时有发生。RNA 病毒的高变异率就决定了 PCR 检测技术和基因芯片检测技术常常难以奏效。随着 HTS 技术的不断改进和测序成本的一再降低,数据分析等相关生物信息学软件的不开发,De Novo Assembly 技术可以帮助我们迅速收集新型病毒的基因组数据,完善病毒基因组数据库,基于 HTS 技术的病毒类 FBPs 检测技术也在迅速发展。

Chiapponi 等^[38]分析了 2 个不同地点的冷冻浆果样本,从这两个样本抽取总 RNA 用于 RNA-seq 和扩增测序,发现两个浆果样本来源的 HAV 基因组序列均与病人体内取得的 HAV 基因组序列具有 99%的同源性。诺如病毒具有极强的传染性,可导致胃肠炎的大范围爆发^[39],也是我国最流行的食源性病毒之一。敖元云等^[40]应用 454 高通量测序技术发现了河北卢龙县存在于腹泻儿童粪便标本中的诺如病毒,经全基因组序列分析发现其与诺如病毒株 30443 同源率为 98.0%。轮状病毒主要感染婴幼儿,是导致婴儿肠胃炎最常见的病原体^[41]。Mamoru 等^[42]运用 Illumina MiSeq sequencing 技术同时对包括诺如病毒、轮状病毒、甲型肝炎病毒和戊型肝炎病毒在内的多种病毒类 FBPs 进行检测,结合微流控技术和 HTS 技术开发出一种高通量检测和基因分型工具。能够有效跟踪多种 RNA 病毒在不同环境中的状态,以更好地了解人类致病性 RNA 病毒。

4.3 真菌和寄生虫类食源性病原体

真菌类 FBPs 主要包括黄曲霉菌、镰刀霉菌、麦角菌

等产毒素真菌,目前,对真菌类 FBPs 的研究远远少于细菌和病毒类 FBPs。真菌类 FBPs 通常首先感染食源性植物,然后在其生长繁殖过程中产生有毒代谢物,人类食用这些植物源性食品而导致疾病^[43]。Vaux 等^[44]利用 Illumina HiSeq 2000 测序平台分析了多例来自于医疗机构的棒地霉感染,结合单核苷酸多态性分析和系统进化分析,发现一种用于储存和输注血液制品的医疗设备为疫情的污染源。

寄生虫类 FBPs 主要包括溶组织内阿米巴虫、绦虫、小隐孢子虫和蓝氏贾第鞭毛虫等。寄生虫类 FBPs 的基因组较大(10~1000 Mbp),是细菌基因组大小的十倍以上,大的甚至接近人类基因组的大小。缺少参考基因组是当前发展基于 HTS 技术的检测寄生虫类 FBPs 的关键阻力,寄生虫基因组数据库正在不断发展更新。Kawai 等^[45]在与多起海产品中毒相关的冷冻牙鲆鱼肌肉组织中找到了可疑孢子,随后直接提取冷冻牙鲆鱼片的总 DNA 和 RNA,分别用于全基因组测序对比分析和 RNA-Seq 分析,发现了导致疫情发生的病原体是粘孢子虫。

到目前为止,将 HTS 技术应用于真菌和寄生虫类 FBPs 的流行病学和毒力研究的报道为数不多。但随着 HTS 技术的不断发展和相关基因组数据库的不断扩充,基于 HTS 技术对真菌和寄生虫类 FBPs 的检测和鉴定将会越来越多地得到应用。

4.4 利用 HTS 技术对食源性病原体日常监测

对一些特定 FBPs 的监测一直是公共卫生领域的常态化工作,目前已有少数国家和实验室已将 HTS 技术列为监测公共健康的常规技术手段,并取得了一定的成果(表 2)。但在全球范围内来说,利用 HTS 技术对食源性病原体日常监测还处于启动阶段。

5 小 结

HTS 技术相比于其它检测技术,可以更加快速、高效、全面地检测食源性病原体,但 HTS 技术的应用需要各种软硬件用于数据的存储和分析,包括数据管理,质量控制、绘图和校准、序列比对组装、基因注释、宏基因组学分析等^[50,51]。然而几乎所有食品科学领域的科学家都缺乏或是只有少量的生物信息学背景,因此大量生物信息学工具的开发和正确使用,并用它们来得到合理的分析结果似乎是一个难以逾越的障碍。因此,需要投入大量努力来培训相关人员使用这些生物信息学工具^[52]。目前,一些已发表的公共的数据分析软件和相关的操作流程已经可以应用于高通量测序数据的查询、下载和可视化的分析^[53,54]。FBPs 基因组数据库的建立是另外一个难点,一个包含所有主要 FBPs 基因组完整信息的参考数据库是流行病学调查和设计病原体检测分析的基础。目前,一些常见的 FBPs 基因组的数据已经释放到公共的数据库^[55-57]。

表2 部分已发表的基于高通量测序技术的食源性病原体检测的研究
Table 2 Researches on food-borne pathogens detection based on HTS technology

食源性病原体	应用的测序技术	数据分析方法	时间地点
细菌类			
大肠杆菌 O157:H7	HiSeq 2500, 200PE	全基因组测序与拼接; 单核苷酸多态分析	2013, 英国 ^[29]
大肠杆菌 O157:H7	HiSeq2000	全基因组测序与拼接; 单核苷酸多态分析	2016, 中国广西 ^[30]
大肠杆菌 O104:H4	MiSeq, 151PE; HiSeq 2500	鉴定环境基因组标签; 基因组比对; 系统进化分析	2011, 德国 ^[46]
单增李斯特菌	MiSeq	全基因组测序与拼接	2017-2018, 南非 ^[32]
肠炎沙门氏菌	MiSeq, 250PE	全基因组测序与拼接; 单核苷酸多态分析	2001-2012, 美国 ^[47]
病毒类			
甲型肝炎病毒	RNA-seq; Sanger amplicon sequencing	与参考基因组比对, 确定基因型	2013, 意大利 ^[38]
诺如病毒	Roche 454	与参考基因组比对; 系统进化分析	2015, 中国河北 ^[40]
诺如病毒	HiSeq	与参考基因组比对; 系统进化分析	2015, 中国广西 ^[48]
轮状病毒	RNA-seq on MiSeq	与参考基因组比对, 确定基因型	2012, 日本 ^[49]
真菌与寄生虫类			
棒地霉	HiSeq2000, 100SE	全基因组测序与拼接, 单核苷酸多态分析	2012, 法国 ^[44]
粘孢子虫	RNA-seq; Sanger amplicon sequencing	基因组拼接与比对; 18S rRNA 序列比对	2008-2010, 日本 ^[45]

HTS 技术可以提高病原体检测的准确性, 特别是对于低丰度的复杂样本, 基于高通量测序技术的 FBPs 检测无需进行选择性地预培养, 通过大量测序数据的分析可以有效防止罕见或未知病原体的漏检。因此, 应用 HTS 技术和宏基因组学分析, 可以快速有效地获知 FBPs 的种属和基因型, 并提供毒力和耐药性等报告。基于高通量测序技术的 FBPs 检测, 是未来 FBPs 检测和日常防控的一条重要思路, 不断完善 FBPs 全基因组数据库, 并开发出更加快速、有效、面向大多数食品检测者的高通量测序数据的分析策略是我们当下需要关注的研究重点。

参考文献

- [1] Marder EP, Cieslak PR, Cronquist AB, *et al.* Incidence and trends of infections with pathogens transmitted commonly through food and the effect of increasing use of culture-independent diagnostic tests on surveillance-foodborne diseases active surveillance network, 10 U.S. Sites, 2013-2016 [J]. *Mmwr Morbid Mort Weekly Report*, 2017, 66(15): 397.
- [2] Roussel A, Charlene A, Sivignon A, *et al.* Foodborne enterotoxigenic *Escherichia coli*: From gut pathogenesis to new preventive strategies involving probiotics [J]. *Future Microbiol*, 2017, 12(1): 73-93.
- [3] Ju J, Xie Y, Guo Y, *et al.* Antibacterial activities of bayberry extract on foodborne pathogens and identification of its active components [J]. *Food Agric Immunol*, 2019, 30(1): 385-397.
- [4] Martinovic T, Andjelic U, Gajdosik MS, *et al.* Foodborne pathogens and their toxins [J]. *Proteomics*, 2016, (147): 226-235.
- [5] Greninger, Alexander L. A decade of RNA virus metagenomics is (not) enough [J]. *Virus Res*, 2017, (244): 218-229.
- [6] Steingart KR, Henry M, Ng V, *et al.* Fluorescence versus conventional sputum smear microscopy for tuberculosis: A systematic review [J]. *Lancet Infect Dis*, 2006, 6(9): 570-581.
- [7] Muniandy S, Dinshaw IJ, Teh SJ, *et al.* Graphene-based label-free electrochemical aptasensor for rapid and sensitive detection of foodborne pathogen [J]. *Anal Bioanal Chem*, 2017, 409(29): 6893-6905.
- [8] Neng J, Li Y, Driscoll AJ, *et al.* Detection of multiple pathogens in serum using silica-encapsulated nanotags in a SERS-based immunoassay [J]. *J Agric Food Chem*, 2018, 66(22): 5707-5712.
- [9] Morgenstern C, Cabric S, Perka C, *et al.* Synovial fluid multiplex PCR is superior to culture for detection of low-virulent pathogens causing periprosthetic joint infection [J]. *Diagn Microbiol Infect Dis*, 2017, 90(2): 115-119.
- [10] Lemieux B, Aharoni A, Schena M. Overview of DNA chip technology [J]. *Mol Breed*, 1998, 4(4): 277-289.
- [11] Sandmann S, De-Graaf AO, Vander-Reijden BA, *et al.* GLM-based optimization of NGS data analysis: A case study of Roche 454, ion torrent PGM and illumina NextSeq sequencing data [J]. *PLoS One*, 2017, 12(2): e0171983.
- [12] Schlaberg R, Chiu CY, Miller S, *et al.* Validation of metagenomic next-Generation sequencing tests for univers [J]. *Arch Pathol Lab Med*, 2017, 141(6): 776-786.
- [13] Frey KG, Herrera-Galeano JE, Redden CL, *et al.* Comparison of three next-generation sequencing platforms for metagenomic sequencing and identification of pathogens in blood [J]. *BMC Genom*, 2014, 15(1): 96.
- [14] Allard MW, Bell R, Ferreira CM, *et al.* Genomics of foodborne pathogens for microbial food safety [J]. *Curr Opin Biotechnol*, 2018, (49): 224-229.
- [15] Ferrario C, Lugli GA, Ossiprandi MC, *et al.* Next generation sequencing-based multigene panel for high throughput detection of food-borne pathogens [J]. *Int J Food Microbiol*, 2017, (256): 20-29.

- [16] Besser J, Carleton HA, Gerner-Smith P, *et al.* Next-generation sequencing technologies and their application to the study and control of bacterial infections [J]. *Clin Microbiol Infect*, 2017, 24(4): 335–341.
- [17] Koduru SV, Tiwari AK, Leberfinger A, *et al.* A comprehensive NGS data analysis of differentially regulated miRNAs, piRNAs, lncRNAs and sn/snoRNAs in triple negative breast cancer [J]. *J Cancer*, 2017, 8(4): 578–596.
- [18] Hwang SG, Kim KH, Lee BM, *et al.* Transcriptome analysis for identifying possible gene regulations during maize root emergence and formation at the initial growth stage [J]. *Genes Genom*, 2018, 40(7): 755–756.
- [19] Ward J, Cole C, Febrer M, *et al.* Almost significant: Simplifying quality control of high-throughput sequencing data [J]. *Bioinformatics*, 2016, 32(24): 3850–3851.
- [20] Li H, Homer N. A survey of sequence alignment algorithms for next-generation sequencing [J]. *Brief Bioinform*, 2010, 11(5): 473–483.
- [21] Edmonson MN, Zhang J, Yan C, *et al.* Bambino: A variant detector and alignment viewer for next-generation sequencing data in the SAM/BAM format [J]. *Bioinformatics*, 2011, 27(6): 865–866.
- [22] Xu X, Sargent DJ, Alan C, *et al.* Genome-wide SNP identification by high-throughput sequencing and selective mapping allows sequence assembly positioning using a framework genetic linkage map [J]. *Bmc Biol*, 2010, 8(1): 1–8.
- [23] Shen X, Shamaiah M, Vikalo H. Iterative learning for reference-Guided DNA Sequence assembly from short reads: Algorithms and limits of performance [J]. *IEEE Transact Signal Process*, 2014, 62(17): 4425–4435.
- [24] Ruffalo M, Laframboise T, Koyuturk M. Comparative analysis of algorithms for next-generation sequencing read alignment [J]. *Bioinformatics*, 2011, 27(20): 2790–2796.
- [25] Stein LD. Using GBrowse 2.0 to visualize and share next-generation sequence data [J]. *Brief Bioinf*, 2013, 14(2): 162–171.
- [26] O'Fallon, Brendan D, Wooderchak-Donahue W, *et al.* A support vector machine for identification of single-nucleotide polymorphisms from next-generation sequencing data [J]. *Bioinformatics*, 2013, 29(11): 1361–1366.
- [27] Zou X, Shi C, Austin RS, *et al.* Genome-wide single nucleotide polymorphism and Insertion-Deletion discovery through next-generation sequencing of reduced representation libraries in common bean [J]. *Mol Breed*, 2014, 33(4): 769–778.
- [28] 肖剑, 陈娟丽, 张彬彬, 等. 大肠杆菌 O157 快速检测板在肉与肉制品检测中的应用[J]. *食品安全质量检测学报*, 2017, (1): 339–343.
Xiao J, Chen LJ, Zhang BB, *et al.* Application of rapid petri-dish in the detection of *Escherichia coli* O157 in meat and meat products [J]. *J Food Saf Qual*, 2017, (1): 339–343.
- [29] Jenkins C, Dallman TJ, Launder N, *et al.* Public health investigation of two outbreaks of shiga toxin-producing *Escherichia coli* O157 associated with consumption of watercress [J]. *Appl Environ Microbiol*, 2015, 81(12): 3946–3952.
- [30] 彭昊, 李军, 杨威, 等. 猪源大肠杆菌 O157:H7 毒力差异株的转录组测序分析[J]. *中国畜牧兽医*, 2017, (6): 1804–1810.
Peng H, Li J, Yang W, *et al.* Transcriptome sequencing analysis of *E. coli* O157:H7 stains of different virulence from pigs [J]. *Chin Anim Husb Vet Med*, 2017, (6): 1804–1810.
- [31] Liu D. Identification, subtyping and virulence determination of *Listeria monocytogenes*, an important foodborne pathogen [J]. *J Med Microbiol*, 2006, 55(6): 645–659.
- [32] Allam M, Tau N, Smouse SL, *et al.* Whole-genome sequences of listeria monocytogenes sequence type 6 isolates associated with a large foodborne outbreak in South Africa, 2017 to 2018 [J]. *Genome Annou*, 2018, 6(25): e00538.
- [33] 吴家林, 沙丹, 马广源, 等. 沙门氏菌 LAMP 检测方法的建立[J]. *中国病原生物学杂志*, 2015, (7): 611–614.
Wu JL, Sha D, Ma GY, *et al.* Establishm entand useofa loop—mediated isothermalamplification techniqueto detectSalmonella [J]. *J Pathogen Biol*, 2015, (7): 611–614.
- [34] Jarvis KG, White JR, Grim CJ, *et al.* Cilantro microbiome before and after nonselective pre-enrichment for *Salmonella* using 16S rRNA and metagenomic sequencing [J]. *BMC Microbiol*, 2015, 15(1): 160.
- [35] RodriguezLázaro D, Cook N, Ruggeri FM, *et al.* Virus hazards from food, water and other contaminated environments [J]. *Fems Microbiol Rev*, 2012, 36(4): 786–814.
- [36] Newell DG, Koopmans M, Verhoef L, *et al.* Food-borne diseases—the challenges of 20 years ago still persist while new ones continue to emerge [J]. *Int J Food Microbiol*, 2010, (139): S3–S15.
- [37] Nieuwenhuijse DF, Koopmans MPG. Metagenomic sequencing for surveillance of food- and waterborne viral diseases [J]. *Front Microbiol*, 2017, (8): 230.
- [38] Chiapponi C, Pavoni E, Bertasi B, *et al.* Isolation and genomic sequence of hepatitis a virus from mixed frozen berries in Italy [J]. *Food EnvironVirol*, 2014, 6(3): 202–206.
- [39] Allen DJ, Harris JP. Methods for ascertaining norovirus disease burdens [J]. *Human Vacc Immunotherapeut*, 2017, 13(11): 2630–2636.
- [40] 敖元云, 虞结梅, 李莉莉, 等. 诺如病毒河北卢龙株 NHBGR59 全基因组序列分析[J]. *中华实验和临床病毒学杂志*, 2016, 30(3): 266–269.
Ao Y, Yu J, Li L, *et al.* Complete sequence analysis of norovirus strain NHBGR59 of Hebei Lulong [J]. *Chin J Exp Clin Virol*, 2016, 30(3): 266–269.
- [41] Gray J, Miren IG. Rotaviruses [J]. *Methods Mol Biol*, 2010, (665): 325–355.
- [42] Mamoru O, Takayuki M, Shinobu K, *et al.* Microfluidic PCR Amplification and MiSeq amplicon sequencing techniques for high-throughput detection and genotyping of human pathogenic RNA Viruses in human feces, sewage, and oysters [J]. *Front Microbiol*, 2018, (9): 830.
- [43] Michael G, Ian P, Sasha T. Ecology and evolution of the human microbiota: Fire, farming and antibiotics [J]. *Genes*, 2015, 6(3): 841–857.
- [44] Vaux S, Criscuolo A, Desnosollivier M, *et al.* Multicenter outbreak of infections by *Saprochaete clavata*, an unrecognized opportunistic fungal pathogen [J]. *Mbio*, 2014, 5(6): e02309.
- [45] Kawai T, Sekizuka T, Yahata Y, *et al.* Identification of Kudoa septempunctata as the causative agent of novel food poisoning outbreaks in Japan by consumption of paralichthys olivaceus in raw fish [J]. *Clin Infect Dis*, 2012, 54(8): 1046–1052.
- [46] Loman NJ, Constantinidou C, Christner M, *et al.* A culture-independent sequence-based metagenomics approach to the investigation of an outbreak of shiga-toxicogenic *Escherichia coli* O104:H4 [J]. *J Am Med Assoc*, 2013, 309(14): 1502–1510.
- [47] Taylor AJ, Lappi V, Wolfgang WJ, *et al.* Characterization of foodborne

- outbreaks of *Salmonella enterica* serovar *Enteritidis* with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection [J]. *J Clin Microbiol*, 2015, 53(10): 3334–3340.
- [48] 信云云, 敖元云, 李利利, 等. 广西猕猴中发现 GII.17 型诺如病毒及其全基因组序列分析[J]. *中华实验和临床病毒学杂志*, 2017, 31(6): 498–503.
- Xin Y, Ao Y, Li L, *et al.* Discovery and gene sequence analysis of GII.17 norovirus in Rhesus monkeys of Guangxi [J]. *Chin J Exp Clin Virol*, 2017, 31(6): 498–503.
- [49] Mizukoshi F, Kuroda M, Tsukagoshi H, *et al.* A food-borne outbreak of gastroenteritis due to genotype G1P[8] rotavirus among adolescents in Japan [J]. *Microbiol Immunol*, 2014, 58(9): 536–539.
- [50] Borges V, Miguel P, Pedro P, *et al.* INSaFLU: An automated open web-based bioinformatics suite “from-reads” for influenza whole-genome-sequencing-based surveillance [J]. *Genome Med*, 2018, 10(1): 46.
- [51] Iacoangeli A, Al-Khleifat A, Sproviero W, *et al.* DNAscan: Personal computer compatible NGS analysis, annotation and visualization [J]. *BMC Bioinform*, 2019, 20(1): 213.
- [52] Taboada EN, Graham MR, Carrico JA, *et al.* Food safety in the age of next generation sequencing, bioinformatics, and open data access [J]. *Front Microbiol*, 2017, (8): 909.
- [53] Egorov AA, Sakharova EA, Anisimova AS, *et al.* Svist4get: A simple visualization tool for genomic tracks from sequencing experiments [J]. *BMC Bioinf*, 2019, 20(1): 113.
- [54] Loh YH, Shen L. Analysis and visualization of ChIP-Seq and RNA-Seq sequence alignments using ngs.plot [J]. *Methods Mol Biol*, 2016, (1415): 371–383.
- [55] Sharma VK, Akavaram S, Schaut RG, *et al.* Comparative genomics reveals structural and functional features specific to the genome of a foodborne *Escherichia coli* O157:H7 [J]. *BMC Genom*, 2019, (20): 196.
- [56] Li Z, Ailyn PO, Wang Y, *et al.* Whole genome sequencing analyses of *Listeria monocytogenes* that persisted in a milkshake machine for a year and caused illnesses in Washington state [J]. *BMC Microbiol*, 2017, 17(1): 134.
- [57] Marton S, Dóro R, Fehér E, *et al.* Whole genome sequencing of a rare rotavirus from archived stool sample demonstrates independent zoonotic origin of human G8P[14] strains in Hungary [J]. *Virus Res*, 2017, (227): 96–103.

(责任编辑: 韩晓红)

作者简介



戴陈伟, 硕士, 助理研究员, 主要研究方向为病原微生物学和分子生物学。
E-mail: 87358066@qq.com



蔡 标, 副研究员, 主要研究方向为病原微生物和流行病学。
E-mail: caib81@163.com