

高通量测序技术在植物及昆虫病毒检测中的应用

战斌慧, 周雪平*

(中国农业科学院植物保护研究所, 植物病虫害生物学国家重点实验室, 北京 100193)

摘要 在过去的十几年中, 测序技术的发展为分子生物学领域带来了革命性的变化。第二代测序(next-generation sequencing, NGS)技术以快速、高灵敏性、高通量、非序列依赖性等特点极大地促进了病毒诊断学研究领域的发展。NGS技术可以在不了解病毒的生物学特性、血清学特点及基因组信息情况下快速检测未知病毒。通过对总核酸样本进行NGS可以获得某个特定生态环境或种植系统中的所有病毒序列, 即病毒组。通过大量的NGS数据可以分析寄主中某一病毒的基因组变化、构建病毒的准种以及研究病毒的进化和起源。本文介绍了高通量测序的方法在植物和昆虫病毒检测中的应用。

关键词 第二代测序; 植物病毒; 昆虫病毒; 检测

中图分类号: S 432.1 **文献标识码:** A **DOI:** 10.16688/j.zwbh.2018326

Application of next-generation sequencing technology in detection of plant and insect viruses

ZHAN Binhui, ZHOU Xueping

(State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing 100193, China)

Abstract In the past decade, the development of sequencing technology has induced revolutionary changes in molecular biology. Next-generation sequencing (NGS) has considerably promoted the development of research on viral diagnostics characterized by rapidity, high sensitivity, high throughput, sequence-independent, etc. NGS made it possible to rapidly detect unknown viruses without knowing the biological characteristics, serological characteristics, and genome information. The whole viral genomes in a particular ecosystem or cropping system, namely virome, could be sequenced through NGS of the total nucleic acids. Viral genome variability, reconstruction of quasispecies, evolution and origin within the host can be obtained by analyzing massive NGS data. In this article, we provide the overview of the process of NGS technology and its applications in detection of plant and insect viruses.

Key words next-generation sequencing; plant virus; insect virus; detection

病毒是一种个体微小、结构简单, 没有细胞结构的特殊生物, 只能在活细胞内寄生并依赖宿主细胞实现增殖。它们种类繁多、数目庞大, 是自然界中最为丰富和多样的生命体^[1-2]。根据宿主的不同, 可以将其分为植物病毒、动物病毒、真菌病毒等。其中, 侵染植物及昆虫的病毒与农业生产密切相关。植物病毒素有植物癌症之称, 其暴发和流行严重危害作物生长, 威胁粮食生产。据不完全统计, 全球每年因植物病毒病造成的损失约占粮食作物总产量的10%^[3-4]。昆虫是农业生态系统的重要组成部分, 其

数目众多, 占已知动物物种的70%以上。已知的昆虫病毒种类有1 000多种, 然而, 目前对昆虫病毒的研究主要集中在对有益昆虫病毒病进行防控以及利用昆虫病毒防治有害昆虫^[5]。因而, 对植物及昆虫病毒进行快速而有效的诊断、发掘新的病毒资源对于增强防控措施、保护农业生产安全具有重要意义。

传统的病毒检测方法主要有: 生物学测定法、电子显微镜观察法^[6]、以酶联免疫吸附反应为代表的血清学检测法^[7]、基于核酸分析的PCR检测法^[8]等。生物学测定法及电子显微镜观察法是最为经典的病

毒检测技术,可用于病毒的初步检测。利用生物学测定法检测植物病毒时,主要依据病毒在指示植物上产生的症状进行诊断,诊断存在主观性;利用电子显微镜观察法可以直观地观察到病毒粒子的存在,但是要求病毒在寄主中具有较高的浓度,而且无法进行种的区分。在过去的几十年中,血清学方法及 PCR 技术为病毒诊断带来了突破性的研究成果,鉴定了大量病毒。然而,上述传统方法的应用均需对病毒的生物学特性、血清学特性、基因组结构、核酸序列等有预先了解,是针对某种或者某类病毒的特异性检测。在针对未知病毒开展的非特异性检测时,以上方法则缺乏优势,检测时间周期长,极大限制了对病毒病害的了解。第二代测序(next-generation sequencing, NGS)技术克服以上传统检测方法的缺点,其所具有的非序列依赖性及高灵敏性为病毒诊断带来了重大的革新^[5,9-10]。本文对高通量测序的方法及其在植物和昆虫病毒检测、资源发掘及进化中的应用作了系统介绍。

1 高通量测序技术

1977 年, Sanger 等利用双脱氧核苷酸末端终止法进行 DNA 测序, 标志着第一代测序技术的诞生^[11-12]。尽管第一代测序技术已经广泛应用于生命科学研究的各个领域, 但由于其成本高、低通量、速度慢等缺点, 无法高效解读所面临的海量生物信息。2005 年, 454 公司(后被 Roche 收购)研发的高通量测序系统 Genome Sequencer 2. 0 System 标志着 NGS 技术的诞生, 使得对遗传信息的分析进入了一个新的时代, 具有里程碑式的意义。主流的测序平台有 Roche 公司的基于焦磷酸测序(pyrosequencing)原理的 454/GS FLX 测序平台、Illumina 公司的基于合成法测序(sequencing by synthesis, SBS)原理的 GAI/HiSeq 平台及 ABI 公司的基于连接法测序(sequencing by ligation, SBL)原理的 SOLiD 平台^[13]。

由于测序原理的不同, 每种测序平台各有其优缺点^[5,9]。Roche 的 454/GS FLX 测序平台具有最长的读长, 可达 700~1 000 bp, 并且运行时间短, 但是通量低, 价格昂贵, 并且在多聚物测量时具有较高的错误率; Illumina 公司的 GAI/HiSeq 平台是目前应用最为广泛的测序平台, 具有通量高、运行速度快、成本低等优点, 具有最高的性价比, 但是在读取复杂模板如高 GC 或高 AT 模板时, 错误率较高; ABI 公司的 SOLiD 平台是目前第二代测序技术中准确率最高的测序平台, 在 15 倍覆盖率时测序准确度可达 99. 999%, 但是读长短, 仅有 75 bp, 且运行时间较长。经过十几年的发展, 各种高通量测序平台也都在不断改进和发展。科研人员可以根据研究的需求不同, 选择相应的测序平台。

2 利用 NGS 检测病毒的流程

除了在生物信息学、生物进化研究、群体遗传学等方面的应用外, NGS 已被应用于病毒的鉴定、诊断及病毒资源的挖掘中。利用 NGS 技术检测病毒和发掘病毒资源的实验流程主要包括: 样品制备、文库构建、平台测序、数据分析以及结果验证^[5,14]。

2.1 样品制备及文库构建

样品的制备是 NGS 的关键步骤, 它将直接影响文库的构建进而影响测序的深度。与寄主基因组相比, 病毒基因组的含量相对较低。所以, 除了利用总 DNA 或总 RNA 制备文库外, 多种有利于提高病毒基因组核酸含量的样品制备方法被应用于病毒检测的实例中, 如测序核酸采用去核糖体 RNA 的总 RNA(ribosomal RNA depleted total RNA, rRNA-depleted totRNA)、双链 RNA(dsRNA)、病毒粒子相关的核酸(virion associated nucleic acids, VANA)、多聚腺苷酸化的 RNA(poly(A) RNA) 及与健康植物抑制消减杂交后的 RNA 等(表 1)。

表 1 用于检测病毒的不同类型核酸样品

Table 1 Different types of nucleic acids for virus detection

核酸类型 Type of nucleic acids	检测的病毒类型 Type of virus	局限 Restrict
总 RNA	ssDNA, dsDNA, ssRNA, dsRNA	检测背景高, 可能会遗漏滴度较低的病毒
总 DNA	ssDNA, dsDNA	不能检测 RNA 病毒
dsRNA	(+) ssRNA, dsRNA	不能检测 (-) ssRNA 和 DNA 病毒
VANA	ssDNA, dsDNA, ssRNA, dsRNA	不能检测无外壳蛋白包被的病毒; 某些病毒粒子提取困难
poly(A) RNA	带有 poly(A) 尾的病毒	遗漏大部分的病毒
rRNA-depleted totRNA	ssDNA, dsDNA, ssRNA, dsRNA	相对高的检测背景
抑制消减杂交后的 RNA	ssDNA, dsDNA, ssRNA, dsRNA	抑制消减杂交文库的构建需要大量时间
sRNA	ssDNA, dsDNA, ssRNA, dsRNA	检测背景高; 对于产生 sRNA 少的病毒, 组装困难大

病毒依据其核酸组成及其复制特征,可以被分为以下几类:双链 DNA 病毒,单链 DNA 病毒,双链 RNA 病毒,正义单链 RNA 病毒,负义单链 RNA 病毒^[15]。在试验过程中应当根据病毒核酸的性质及其在侵染过程中的特点,选择不同的核酸样品进行文库制备和 NGS 测定。

dsRNA 是 RNA 病毒在侵染循环过程中所产生的特有的形态,因此,利用 dsRNA 作为测序核酸可以高效地获得病毒的特异序列,提高测序的深度。Al-Rwahnih 等^[16]利用 NGS 检测引起西拉葡萄衰退症状的病毒时,分别提取了总 RNA 和 dsRNA 制备文库。测序结果表明,利用 dsRNA 构建的文库中病毒来源的 reads 数占总 reads 数的 52.7%(54 605/103 597),而利用总 RNA 构建的文库中病毒来源的 reads 数仅占 1.94%(1 275/65 587),说明通过提取 dsRNA 富集了较多的病毒 dsRNA,提高了检测的灵敏性。数据分析发现共有沙地葡萄茎痘伴随病毒 *Grapevine rupestris stem pitting-associated virus* (GRSPaV)、沙地葡萄叶脉羽化病毒 *Grapevine rupestris vein-feathering virus* (GRVfV) 等 6 个已知病毒及类病毒和一个新病毒西拉葡萄病毒 1 号 *Grapevine syrah virus-1* (GSyV-1) 侵染西拉葡萄。另外,利用 dsRNA 作为检测病毒的核酸样品是检测真菌病毒的最主要方法。利用此方法可以有效地检测 RNA 病毒,但却不能有效地检测 DNA 病毒^[17]。另外,双链 RNA 的提取花费时间久,操作复杂。

利用 VANA 进行 NGS 测序是更为直接有效鉴定病毒的方法。近年来,利用此策略已经成功测定了多种植物及昆虫病毒。例如,Candresse 等^[18]利用 454 测序平台测定了命名为 VARX 和 VSDA 的两个甘蔗样品的 VANA,除了已知的埃及甘蔗线条病毒 *Sugarcane streak Egypt virus* (SSEV) 外,还发现一种 DNA 病毒,命名为甘蔗白线条病毒 *Sugarcane white streak virus* (SWSV)。数据分析表明,来源于 SSEV 的 reads 数在 VARX 和 VSDA 样品中分别占 53.7%(1 398/2 612) 和 75%(1 227/1 635),覆盖整个 SSEV 基因组 159 倍和 138 倍;来源于 SWSV 的 reads 数在 VARX 和 VSDA 样品中分别占 23.9%(625/2 612) 和 16.1%(264/1 635),覆盖整个基因组 81 倍和 29 倍。表明通过纯化 VANA 能够有效去除基因组的序列,达到最大化富集病毒核酸的目的。但是,对于无外壳蛋白包裹的

病毒此方法则无法制备样品。另外,不同科属的病毒粒子提取方法差异较大,当样本中可能包含两种或几种病毒时,不能用同一种方法同时提取所有的病毒粒子。因而,在对检测样本了解较少的前提下,样品制备困难较大。

利用富集 poly(A)RNA 的方法制备样本,可以用于检测 DNA 和 RNA 病毒,但是不能检测没有 poly(A)结构的病毒。Wylie 和 Jones^[19]提取分离了水仙和朱顶红样品的 poly(A)RNA,并以此为模板构建文库测序,检测到了麝香石竹潜隐病毒属和马铃薯 Y 病毒属的 7 个不同的病毒分离物。Wylie 等^[20]以相似的方法检测到分离自澳大利亚的百香果木质病毒 *Passion fruit woodiness virus* (PWV),病毒的 RNA 占分离的 poly(A)RNA 的 7.38%。

在植物病毒鉴定中,抑制消减杂交技术被用于样品的制备,Adams 等^[21]利用未感染病毒的 cDNA 对感病植物的 cDNA 进行消减杂交,并利用 NGS 鉴定到了黄瓜花叶病毒属的新病毒,名为蛇鞭菊轻斑驳病毒 *Gayfeather mild mottle virus* (GMMV)。Monger 等^[22-23]也利用构建抑制消减文库的方式检测到了木薯褐条病毒 *Cassava brown streak virus* (CBSV) 和美人蕉黄条病毒 *Canna yellow streak virus* (CaYSV)。但是,抑制消减文库的构建需要花费大量的时间,因而并不适合高通量样品的检测和病害诊断。

目前,总 RNA 测序适用于多种类型的病毒基因组,均能够检测 DNA 病毒及 RNA 病毒,并且样品制备相对简单,可被大部分的检测实验室使用。据统计,仅在 2009 至 2014 年间,就有 30 例利用总 RNA 构建的文库进行 NGS 测序检测植物病毒的实例被报道^[17],然而在检测较低滴度的病毒时有很大的不足。去除总 RNA 中含量丰富的 rRNA,增加病毒 RNA 的相对含量则克服了以上缺点。研究表明,以 rRNA-depleted totRNA 构建文库能够使病毒 RNA 的相对含量提高 10 倍^[24]。Liu 等^[25]利用感病的小麦和带毒的叶蝉样品分别构建 rRNA-depleted totRNA 文库,通过测序检测到一种新的经叶蝉传播的弹状病毒,名为小麦黄条纹病毒 *Wheat yellow striate virus* (WYSV)。Xin 等^[26]通过分析感病西瓜样品的 rRNA-depleted totRNA 文库,检测到两种新的负义单链 RNA 病毒,并建议归属为布尼亚病毒目 *Bunyavirales* 白纤病毒科 *Phenuiviridae*。

另外,小 RNA(sRNA)测序也是鉴定病毒最常用

的方法之一。自 2009 年,Donaire 等^[27]和 Kreuze 等^[28]证明可以利用病毒来源的 sRNA 进行病毒基因组的组装以鉴定植物中的已知和未知病毒后,这一方法被广泛应用于未知病毒的鉴定。依据的原理是病毒的侵染会诱导寄主的 RNA 沉默机制发挥作用,产生病毒来源的 sRNA。通过对寄主中 sRNA 的测序及分析可以鉴定寄主中的病毒种类。该方法应用范围广,可以检测不同类型的 DNA 及 RNA 病毒^[29]。迄今,应用该方法已成功检测 30 多种植物病毒^[17]。但对于某些产生 sRNA 较少的持久性病毒或者潜伏性病毒,可能不能精确地组装病毒基因组。

2.2 平台检测及数据分析

在检测某种植物或昆虫病毒时,可根据侧重检测的病毒种类或已有的提示,提取或富集相应的核酸样本,根据检测需要,选取相应的测序平台获得高通量的测序数据。数据分析是利用 NGS 检测病毒的关键点和难点。以下以分析总 RNA 或 rRNA-depleted totRNA 测序数据为例,说明数据分析的过程^[5]。第一步,原始数据处理,运用 EMBOSS、CLC Genomic Workbench 等软件去掉两端的接头、去掉低质量的数据等,获得 clean reads。第二步,去掉基因组的数据(可选),若已知寄主的基因组数据库,可利用 CLC Genomic Workbench 等软件将获得的 clean reads 与参考基因组比对,筛选出不能比对到参照基因组的序列做进一步分析。第三步,将筛选的序列利用 CLC Genomic Workbench 进行从头拼接,然后将获得的片段(contigs)进行 BLAST 分析。第四步,对具有同源性的片段进一步组装,并比对 NCBI 核酸及蛋白数据库,揭示病毒的种类。

2.3 结果验证

根据获得的 contigs,提取寄主的核酸,运用 PCR、反转录 PCR、cDNA 末端快速扩增(rapid amplification of cDNA ends, RACE)等技术验证病毒并扩增病毒基因组的全长序列。

3 NGS 在植物及昆虫病毒检测及发掘中的应用

3.1 疑难杂症的病原诊断及新病毒的发现

基于 NGS 可应用于非特异性检测病毒的特点,很多农业生产中的疑难杂症可以利用 NGS 进行诊断。桑花叶型萎缩病是经济作物桑树上的重要病害,感病的植株较健康植株矮小,叶片花叶、卷曲,严重影响桑树的经济价值^[30]。在长达数十年中,此病

的病原未能明确,因而不能得到有效的防治。通过 sRNA 测序,从感病植株中检测到一种新的 DNA 病毒,研究表明该病毒与桑花叶型萎缩病具有极大地相关性,命名为桑花叶萎缩相关病毒 Mulberry mosaic dwarf associated virus (MMDaV)^[31]。我国是西瓜的主要生产国,种植面积及产量均居世界首位。病毒病害是限制西瓜产量的重要因素之一。2015—2016 年,Xin 等^[26]在对河南开封的西瓜病害开展的田间调查中发现了表现花叶、卷曲症状的西瓜植株。利用总 RNA 测序及 sRNA 测序检测引起此病害的病毒,鉴定出了两种负义单链 RNA 病毒,命名为西瓜皱叶病毒 1 号 Watermelon crinkle leaf-associated 1 (WCLaV-1)和西瓜皱叶病毒 2 号 Watermelon crinkle leaf-associated 2 (WCLaV-2)。进一步的接种试验表明 WCLaV-1 能够引起与田间西瓜相似的皱叶症状。2013—2014 年,Chen 等^[32]在云南和贵州省玉米病毒病的调查中发现表现典型的黄化、花叶症状的玉米植株。利用 sRNA 测序鉴定引起此病害的病毒,发现一种新的病毒,命名为玉米黄化花叶病毒 Maize yellow mosaic virus (MaYMV)。2006—2007 年,蜜蜂的蜂群崩溃失调病(colony collapse disorder, CCD)席卷美国的 22 个州,同时在法国、德国、澳大利亚等国也引起严重危害^[33]。CCD 的主要症状是大量的成年工蜂在短期内突然消失在蜂巢外,没有任何尸体,只有蜂王、卵、未成年工蜂残存在蜂巢中。为了探究 CCD 的病因,Cox-Foster 等^[34]来自美国患 CCD 的蜂群和来自美国及澳大利亚的健康蜂群中提取总 RNA 构建文库测序,生物信息学分析从患 CCD 的蜂群中发现了 7 种病毒而从健康蜂群中发现了 5 种病毒。通过比对发现,以色列急性麻痹病毒 Israeli acute paralysis virus (IAPV)与克什米尔蜜蜂病毒 Kashmir bee virus (KBV)可能与 CCD 有关,通过进一步的研究确认 IAPV 与 CCD 的相关性最大。

近年来,通过 NGS 技术鉴定了很多不引起明显症状的潜伏性病毒,极大地丰富了病毒的资源。混合病毒科 *Amalgaviridae* 是由国际病毒分类委员会于 2013 年设立的一个新科,它由一类 dsRNA 病毒组成。因兼具整体病毒科 *Totiviridae* 和分体病毒科 *Partitiviridae* 的特征,故称为混合病毒科。此科包含 4 个确定种,其病毒粒子在侵染的植物组织中很难被纯化和观察到,并且并不能确定它们与寄主植物的某种症状具有相关性^[35]。因而,发现此科病

毒的侵染以及鉴定此科的病毒均具有很大的难度。然而,近两年由于生物信息学的快速发展,通过高通量的数据分析,发现和鉴定了很多之前未发现的病毒,极大地丰富了此科病毒的种类。例如, Nibert 等^[36]通过分析转录组数据,从 12 种不同的植物中筛选到 19 个包含混合病毒科病毒的序列。其中的 16 段序列包含预测混合病毒的全长编码区。Park 等通过分析大叶藻^[37]和菠菜^[38]的转录组数据,鉴定了两种侵染大叶藻及一种侵染菠菜的混合病毒科病毒。

3.2 病毒组的鉴定

宏基因组学(metagenomics)是以测序为手段分析某个特定生态环境中整个微生物群体的基因组。在病毒学研究领域,基于 NGS 的宏基因组学分析使得研究某个特定生态系统或种植系统中的所有病毒成为可能。整个病毒的群体被称为病毒组(virome)。2009 年, Roossinck 等^[39]发起了一个研究病毒生物多样性的调查。他们收集了 10~20 g 植物材料用于宏基因组测序。这些植物材料采集于美国俄克拉荷马州东北部的高草草原保护区(具有较低的植物多样性)和哥斯达黎加西北部关纳卡斯帝保护区(具有较高的植物多样性),涉及 15 个科 400 余个样品。生物信息学分析表明高达 70% 的样品中有病毒的同源序列,包括雀麦花叶病毒科、花椰菜花叶病毒科、分体病毒科、马铃薯 Y 病毒科等 11 个科的病毒,还包括一些未分类的病毒。2012 年, Wylie 等^[40]从 17 种植物 120 个不同的植物叶片中提取总 RNA 并分离 poly(A)RNA 用于 NGS 分析。从中鉴定到 20 种带 poly(A)RNA 尾的 RNA 病毒,其中 16 个为已知病毒,4 个为新病毒。2011 年, Coetzee 等^[41]从南非的一个葡萄园中随机选取了 44 个葡萄藤蔓,从中分离 dsRNA,构建文库并进行 NGS。利用生物信息学方法构建病毒群体,包含 4 种已知的病毒及几种新病毒,其中葡萄卷叶伴随病毒 3 号 Grapevine leafroll-associated virus 3 (GLRaV-3)为优势病毒,包含该病毒的 reads 数占总 reads 数的 59%。

番茄是一种重要的经济作物,病毒病是限制番茄生产的主要因素。我国是番茄的主要生产国,为了研究番茄病毒在我国的分布特点,为番茄病毒病防治提供线索, Xu 等^[42]从我国主要番茄栽培区采集了 170 个疑似病毒侵染的番茄样品。通过 sRNA 测序及数据分析获得侵染我国番茄的病毒组,共检测到 21 种已知病毒、一个新发现的番茄病毒以及两

种类病毒。病毒种类涉及 12 个属,其中正义单链 RNA 病毒是主要的群体。89% 的样品为混合侵染,包含两种或三种病毒。通过分析病毒组还发现,在我国主要侵染番茄的病毒有番茄花叶病毒 Tomato mosaic virus (ToMV)、番茄黄化曲叶病毒 Tomato yellow leaf curl virus (TYLCV)、马铃薯 Y 病毒 Potato virus Y (PVY)、南方番茄病毒 Southern tomato virus (STV)等 10 种病毒。分析番茄病毒组可以检测到重组事件的发生率,黄瓜花叶病毒 Cucumber mosaic virus (CMV)和 ToMV 在正义-正义链重组和正义-负义链重组中均具有较高的重组率。另外,通过病毒组的分析还可以预测病毒的起源及进化等,利用各个地区 TYLCV 不同株系的基因组序列构建进化树,发现所有已经报道的来自中国的 TYLCV 株系在进化树中归于一支,表明其具有共同的起源;筛选 56 个代表株系,利用贝叶斯系统发育法分析其进化率,推测 TYLCV 可能早在 1996 年就已经在我国发生,早于 2006 年的首次报道。

3.3 病毒准种的构建

由于较高的突变和重组效率,病毒(特别是 RNA 病毒)是进化最快的生物之一。某一寄主中的病毒群体并不是单一的序列,而是包含很多不完全一致但是相似度极高的病毒基因组序列,称为病毒的准种(quasispecies)^[43]。了解准种的遗传信息对于了解病毒的变异、进化、传播、毒性、躲避寄主的防御反应等均具有重要的意义。利用 NGS 构建病毒的准种、了解其群体变异主要集中在与人类疾病相关的病毒研究中,如登革病毒^[44]、手足口病毒^[45]、乙型肝炎病毒^[46]等。近年来,利用 NGS 对植物和昆虫病毒准种的研究也取得了一些进展。Fabre 等^[47]利用 NGS 技术分析了 4 种具有不同替换位点及致病性特点的 PVY 变体在感病寄主辣椒中的种群动态,并结合数学模型描述了自然选择和遗传漂变在改变病毒的进化动力学中具有重要作用。Cornman 等^[48]通过 NGS 研究了侵染蜜蜂的蜜蜂畸翅病毒和 IAPV 的群体遗传变异,阐明不同区域不同样品间的单核苷酸多态性谱与侵染滴度和寄主种群数目具有相关性。Huang 等^[49]通过 NGS 分析和比对水稻条纹病毒 Rice stripe virus (RSV)在自然寄主水稻和试验寄主本氏烟中的遗传多样性,发现在本氏烟中共有 1 675 处位点发生替换而在水稻中仅有 341 处,表明 RSV 在本氏烟中具有更高的遗传多样性,

推测与寄主适应性相关。

4 优势和局限性

与传统的病毒检测方法相比,NGS 的优势主要表现在以下几个方面。(a)NGS 检测具有非序列依赖性,不需要依赖于病毒的序列特征或者生物学特征^[14]。分析获得的 NGS 数据时,可以通过序列同源性或者非序列同源性两种方式来分析潜在的病毒序列。序列同源性的方式是基于未知病毒可能与已报道的病毒序列具有一定程度的同源性,进而推测获得的 contigs 可能包含与已知病毒具有同源性的新病毒^[5];非序列同源性的方式则是采用逐渐过滤重叠 sRNA 的方式,发现与已知病毒序列完全不具有同源性的新病毒^[50]。(b)NGS 检测具有快速、灵敏性高的特点。利用传统检测方法检测未知病毒时,由于对病毒的序列及生物学特性了解甚少,所以检测花费的时间会非常长,甚至不能取得理想的检测结果。而 NGS 一般在几天内即可完成测序反应,数据分析及结果验证也可在较短时间内完成^[51]。(c)NGS 可以同时检测多种不同核酸类型的病毒。不论是 DNA 病毒或者 RNA 病毒均能够在一次总 RNA 或 sRNA 测序反应中被检测到^[17]。实际应用中也可提取不同类型的核酸构建文库,以提高测序的准确性、灵敏性及广泛性^[52]。(d)NGS 可同时检测多个样本,样品检测具有高通量的特点。可将每个样品的 cDNA 带上特异性的序列,这样可同时检测来自于不同地区的不同种寄主中病毒的种类、分布以及变异进化等特征^[39]。

NGS 的局限性主要包括:(a)相较于 Sanger 测序而言,NGS 的错误率相对较高且读长较短。(b)NGS 需要科研人员掌握较高的生物信息学技术。NGS 会产生海量的数据,如何对这些遗传信息进行筛选和分析,需要研究人员精通生物信息学分析方法。(c)目前 NGS 不能获得全长的基因组序列,特别是在 5' 端和 3' 端的非编码区,因而需要结合 RACE-PCR 等方法才能获得完整的病毒序列^[5]。对于某些滴度比较低的病毒,在组装后仅能获得较短的 contigs,需要与其他分子生物学试验结合才能确定病毒的存在。(d)通过 NGS 能够获得病毒的遗传信息,但是还需要掌握病毒的传播方式、寄主范围等生物学特性才能对新病毒进行分类以及证实病毒与某种疾病之间的相关性^[14,51]。

5 结语

NGS 的诞生为病毒学研究带来了新的方法,极大地推动了病毒学的发展。植物病毒和昆虫病毒与农业生产紧密相连,未来随着 NGS 的成本降低以及广泛推广,其将在病毒病快速诊断、病毒资源发掘、病毒组分析、准种构建等研究领域取得更为显著的成果,为农业的健康生产提供理论指导。以 PacBio 为代表的三代测序平台能够直接针对单个分子进行测序,具有准确率高、读长长、无 GC 偏好性等特点,但目前测序费用昂贵未能被广泛推广,期望未来该技术能够促进病毒学的发展。

参考文献

- [1] SUTTLE C A. Marine viruses-major players in the global ecosystem [J]. *Nature Reviews Microbiology*, 2007, 5(10): 801 - 812.
- [2] SUTTLE C. The virosphere: the greatest biological diversity on earth and driver of global processes [J]. *Environmental Microbiology*, 2005, 7(4): 481 - 482.
- [3] STRANGE R N, SCOTT P. Plant disease: a threat to global food security [J]. *Annual Review of Phytopathology*, 2005, 43(1): 83 - 116.
- [4] COUTTS R. Plant viruses: new threats to crops from changing farming practices [J]. *Nature*, 1986, 322(6080): 594.
- [5] LIU Sijun, VIJAYENDRAN D, BONNING B C. Next generation sequencing technologies for insect virus discovery [J]. *Viruses*, 2011, 3(10): 1849 - 1869.
- [6] GOLDSMITH C S, MILLER S E. Modern uses of electron microscopy for detection of viruses [J]. *Clinical Microbiology Reviews*, 2009, 22(4): 552 - 563.
- [7] KOENIG R. Indirect ELISA methods for the broad specificity detection of plant viruses [J]. *Journal of General Virology*, 1981, 55(1): 53 - 62.
- [8] ORTIZ V, CASTRO S, ROMERO J. Optimization of RT-PCR for the detection of bean leaf roll virus in plant hosts and insect vectors [J]. *Journal of Phytopathology*, 2010, 153(2): 68 - 72.
- [9] BARBA M, CZOSNEK H, HADIDI A. Historical perspective, development and applications of next-generation sequencing in plant virology [J]. *Viruses*, 2014, 6: 106 - 136.
- [10] HADIDI A, FLORES R, CANDRESSE T, et al. Next-generation sequencing and genome editing in plant virology [J]. *Frontiers in Microbiology*, 2016, 7: 1325.
- [11] SANGER F, NICKLEN S, COULSON A R. DNA sequencing with chain-terminating inhibitors [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1977, 74: 5463 - 5467.
- [12] SANGER F, AIR G M, BARRELL B G, et al. Nucleotide sequence of bacteriophage ϕ x174 DNA [J]. *Nature*, 1977, 265(5596): 687 - 695.

- [13] LIU Lin, LI Yihu, LI Siliang, et al. Comparison of next-generation sequencing systems [J]. *Journal of Biomedicine & Biotechnology*, 2012, 2012; 251364.
- [14] 钱亚娟, 徐毅, 周琦, 等. 利用深度测序技术发掘植物病毒资源 [J]. *中国科学: 生命科学*, 2014, 44(4): 351 - 363.
- [15] 赫尔. 马修斯植物病毒学 [M]. 范在丰, 李怀芳, 韩成贵, 等, 译. 北京: 科学出版社, 2007: 190 - 207.
- [16] AL-RWAHNIH M, DAUBERT S, GOLINO D, et al. Deep sequencing analysis of RNAs from a grapevine showing syrah decline symptoms reveals a multiple virus infection that includes a novel virus [J]. *Virology*, 2009, 387(2): 395 - 401.
- [17] ROOSSINCK M, MARTIN D P, ROUMAGNAC P. Plant virus metagenomics: Advances in virus discovery [J]. *Phytopathology*, 2015, 105(6): 716 - 727.
- [18] CANDRESSE T, FILLOUX D, MUHIRE B, et al. Appearances can be deceptive; revealing a hidden viral infection with deep sequencing in a plant quarantine context [J/OL]. *PLoS ONE*, 2014, 9: e102945.
- [19] WYLIE S J, JONES M G K. Complete genome sequences of seven carlavirus and potyvirus isolates from *Narcissus* and *Hippeastrum* plants in Australia, and proposals to clarify their naming [J]. *Archives of Virology*, 2012, 157(8): 1471 - 1480.
- [20] WYLIE S J, JONES M G K. The complete genome sequence of a passion fruit woodiness virus isolate from Australia determined using deep sequencing, and its relationship to other potyviruses [J]. *Archives of Virology*, 2011, 156(3): 479 - 482.
- [21] ADAMS I P, GLOVER R H, MONGER W A, et al. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology [J]. *Molecular Plant Pathology*, 2009, 10(4): 537 - 545.
- [22] MONGER W A, ALICAI T, NDUNGURU J, et al. The complete genome sequence of the Tanzanian strain of Cassava brown streak virus and comparison with the Ugandan strain sequence [J]. *Archives of Virology*, 2010, 155(3): 429 - 433.
- [23] MONGER W A, ADAMS I P, GLOVER R H, et al. The complete genome sequence of Canna yellow streak virus [J]. *Archives of Virology*, 2010, 155(9): 1515 - 1518.
- [24] ADAMS I, FOX A. Diagnosis of plant viruses using next-generation sequencing and metagenomic analysis [M]// WANG A, ZHOU X. *Current research topics in plant virology*. Switzerland: Springer Cham, 2016: 323 - 335.
- [25] LIU Yan, DU Zhenzhen, WANG Hui, et al. Identification and characterization of wheat yellow striate virus, a novel leafhopper-transmitted nucleorhabdovirus infecting wheat [J]. *Frontiers in Microbiology*, 2018, 9: 468.
- [26] XIN Min, CAO Mengji, LIU Wenwen, et al. Two negative-strand RNA viruses identified in watermelon represent a novel clade in the order *Bunyavirales* [J]. *Frontiers in Microbiology*, 2017, 8: 1514.
- [27] DONAIRE L, WANG Yu, GONZALEZ-IBEAS D, et al. Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes [J]. *Virology*, 2009, 392(2): 203 - 214.
- [28] KREUZE J F, PEREZ A, UNTIVEROS M, et al. Complete viral genome sequence and discovery of novel viruses by deep sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of viruses [J]. *Virology*, 2009, 388(1): 1 - 7.
- [29] PECMAN A, KUTNJAK D, GUTIÉRREZ-AGUIRRE I, et al. Next generation sequencing for detection and discovery of plant viruses and viroids: comparison of two approaches [J]. *Frontiers in Microbiology*, 2017, 8: 1998.
- [30] 薛松, 陶成. 桑树花叶卷叶病的发生及防治 [J]. *中国蚕业*, 2003, 24(1): 35.
- [31] MA Yuxin, NAVARRO B, ZHANG Zhixiang, et al. Identification and molecular characterization of a novel monopartite geminivirus associated with mulberry mosaic dwarf disease [J]. *Journal of General Virology*, 2015, 96(8): 2421 - 2434.
- [32] CHEN Sha, JIANG Guangzhuang, WU Jianxiang, et al. Characterization of a novel polerovirus infecting maize in China [J]. *Viruses*, 2016, 8: 120.
- [33] VAN ENGELSDORP D, MEIXNER M D. A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them [J]. *Journal of Invertebrate Pathology*, 2010, 103: S80-S95.
- [34] COX-FOSTER D L, CONLAN S, HOLMES E C, et al. A metagenomic survey of microbes in honey bee colony collapse disorder [J]. *Science*, 2007, 318(5848): 283 - 287.
- [35] KRUPOVIC M, DOLJA V V, KOONIN E V. Plant viruses of the *Amalgaviridae* family evolved via recombination between viruses with double-stranded and negative-strand RNA genomes [J]. *Biology Direct*, 2015, 10: 12.
- [36] NIBERT M L, PYLE J D, FIRTH A E. A+1 ribosomal frameshifting motif prevalent among plant amalgaviruses [J]. *Virology*, 2016, 498: 201 - 208.
- [37] PARK D, GOH C J, KIM H, et al. Identification of two novel amalgaviruses in the common eelgrass (*Zostera marina*) and in silico analysis of the amalgavirus +1 programmed ribosomal frameshifting sites [J]. *Plant Pathology Journal*, 2018, 34(2): 150 - 156.
- [38] PARK D, HAHN Y. Genome sequences of spinach deltapartitivirus 1, spinach amalgavirus 1, and spinach latent virus identified in spinach transcriptome [J]. *Journal of Microbiology & Biotechnology*, 2017, 27(7): 1324 - 1330.
- [39] ROOSSINCK M J, SAHA P, WILEY G B, et al. Ecogenomics: using massively parallel pyrosequencing to understand virus ecology [J]. *Molecular Ecology*, 2010, 19(S1): 81 - 88.
- [40] WYLIE S J, LUO Hao, LI Hua, et al. Multiple polyadenylated RNA viruses detected in pooled cultivated and wild plant samples [J]. *Archives of Virology*, 2012, 157(2): 271 - 284.
- [41] COETZEE B, FREEBOROUGH M J, MAREE H J, et al. Deep sequencing analysis of viruses infecting grapevines: virome of a vineyard [J]. *Virology*, 2010, 400(2): 157 - 163.
- [42] XU Chenxi, SUN Xuepeng, TAYLOR A, et al. Diversity, distribution, and evolution of tomato viruses in china uncovered by small RNA sequencing [J/OL]. *Journal of Virology*, 2017, 91(11): e00173 - 17.

- [24] 刘万才,刘宇,曾娟,等. 推进农业有害生物数字化监测预警建设刍议[J]. 中国植保导刊,2009,29(10):11-15.
- [25] 刘万才,刘宇,龚一飞. 论重大有害生物数字化监测预警建设的长期任务[J]. 中国植保导刊,2011,31(1):25-29.
- [26] 刘宇,刘万才,韩梅. 农作物重大病虫害数字化监测预警系统建设进展[J]. 中国植保导刊,2011,31(2):33-35.
- [27] 刘宇,刘万才,王学锋. 水稻重大病虫害数字化监测预警平台的设计与实现[J]. 中国植保导刊,2009,29(12):5-9.
- [28] 郑庆伟. 新疆农业有害生物远程预警防控指挥中心正式启动[J]. 农药市场信息,2016(5):61.
- [29] 曹烁,王睿文,王鹏,等. 河北省植物保护网络服务平台的实践与思考[J]. 中国植保导刊,2016,36(10):78-80.
- [30] 陈海中,张友华,刘家成,等. 安徽省农作物病虫监测预警平台的研制[J]. 中国植保导刊,2013,33(11):54-58.
- [31] 刘万才,姜玉英,曾娟,等. 电视—手机—网络三位一体现代病虫预报发布新模式的创新与应用[J]. 中国植保导刊,2013,33(6):47-49.
- [32] 龚一飞,刘万才. 病虫预报网络发布方式的创新与实践[J]. 中国植保导刊,2011,31(11):48-50.
- [33] 黄冲,刘万才,姜玉英,等. 微信公众号发布病虫情报的创新与实践[J]. 中国植保导刊,2017,37(10):30-34.
- [34] 姜玉英,刘万才,曾娟,等. 农作物病虫测报工作年报 2010[M]. 北京:中国农业出版社,2011.
- [35] 姜玉英,刘万才,曾娟,等. 农作物病虫测报工作年报 2011[M]. 北京:中国农业出版社,2012.
- [36] 曾娟,黄冲,刘万才,等. 农作物病虫测报工作年报 2012[M]. 北京:中国农业出版社,2013.
- [37] 黄冲,陆明红,刘万才,等. 农作物病虫测报工作年报 2013[M]. 北京:中国农业出版社,2014.
- [38] 陆明红,刘杰,刘万才,等. 农作物病虫测报工作年报 2014[M]. 北京:中国农业出版社,2015.
- [39] 刘万才,刘杰,姜玉英,等. 农作物病虫测报工作年报 2015[M]. 北京:中国农业出版社,2016.
- [40] 刘万才,姜玉英,龚一飞,等. 病虫测报经验与启示[M]. 北京:中国农业出版社,2010.
- [41] 刘万才,姜玉英,陆明红,等. 病虫测报创新与实践[M]. 北京:中国农业出版社,2012.
- [42] 刘万才,姜玉英,黄冲,等. 病虫测报数字化与信息化[M]. 北京:中国农业出版社,2013.
- [43] 刘万才,黄冲,刘杰. 韩国农作物有害生物监测预警建设的经验[J]. 世界农业,2016(5):59-63.
- [44] 刘万才,陆明红,翟保平,等. 越南水稻生产及其迁飞性害虫发生情况[J]. 中国植保导刊,2014,34(10):91-95.
- [45] 中华人民共和国农业部. 农业部关于印发《到 2020 年化肥使用量零增长行动方案》和《到 2020 年农药使用量零增长行动方案》的通知[EB/OL]. (2016-01-15)[2017-05-14]. <https://wenku.baidu.com/view/7c83c9c614791711cd791764.html?from=search###>.
- [46] 中共中央,国务院. 关于深入推进农业供给侧结构性改革,加快培育农业农村发展新动能的若干意见[EB/OL]. (2017-02-09)[2017-05-14]. <http://hh.hljagri.gov.cn/detail/7686.html>.
- [47] 刘万才,黄冲,陆明红,等. 近 10 年我国农作物主要病虫害发生危害情况的统计和分析[J]. 植物保护,2016,42(5):1-9.
- [48] 农业部农作物总站. 农作物主要病虫害测报办法[M]. 北京:农业出版社,1981:1-290.
- [49] 全国农业技术推广服务中心. 主要农作物病虫害测报技术规范应用手册[M]. 北京:中国农业出版社,2010:1-292.

(责任编辑:田喆)

(上接 126 页)

- [43] LAURING A S, ANDINO R. Quasispecies theory and the behavior of RNA viruses [J/OL]. PLoS Pathogens, 2010, 6: e1001005.
- [44] CHININMANU K, SUT'TITHEPTUMRONG A, SANGSRAKRU D, et al. Feasibility of using 454 pyrosequencing for studying quasispecies of the whole dengue viral genome [J]. BMC Genomics, 2012, 13: S7.
- [45] WRIGHT C F, MORELLI M J, THÉBAUD G, et al. Beyond the consensus: dissecting within-host viral population diversity of foot-and-mouth disease virus by using next-generation genome sequencing [J]. Journal of Virology, 2011, 85(5): 2266-2275.
- [46] LI F, ZHANG D, LI Y, et al. Whole genome characterization of hepatitis B virus quasispecies with massively parallel pyrosequencing [J]. Clinical Microbiology and Infection, 2015, 21(3): 280-287.
- [47] FABRE F, MONTARRY J, COVILLE J, et al. Modelling the evolutionary dynamics of viruses within their hosts: a case study using high-throughput sequencing [J/OL]. PLoS Pathogens, 2012, 8: e1002654.
- [48] CORNMAN R S, BONCRISTIANI H, DAINAT B, et al. Population-genomic variation within RNA viruses of the western honey bee, *Apis mellifera*, inferred from deep sequencing [J]. BMC Genomics, 2013, 14: 154.
- [49] HUANG Lingzhe, LI Zefeng, WU Jianxiang, et al. Analysis of genetic variation and diversity of rice stripe virus populations through high-throughput sequencing [J]. Frontiers in Plant Science, 2015, 6: 176.
- [50] WU Qingfa, WANG Ying, CAO Mengji, et al. Homology-independent discovery of replicating pathogenic circular RNAs by deep sequencing and a new computational algorithm [J]. Proceedings of the National Academy of Sciences of the United States of America, 2012, 109(10): 3938-3943.
- [51] 马宇欣,李世访. 高通量测序技术在鉴定木本植物双生病毒中的应用[J]. 植物保护, 2016, 42(6): 1-10.
- [52] LOCONSOLE G, GIAMPETRUZZI A, ROBERTO R, et al. Next-generation sequencing and metagenomic analysis advances in plant virus diagnosis and discovery [J]. Journal of Plant Pathology, 2012, 94: 48-49.

(责任编辑:杨明丽)