

基于小 RNA 深度测序技术首次发现 侵染黄槿的杆状 DNA 病毒

王致远^{1#}, 冯如盈^{1#}, 娄义念¹, 梁凯莉¹, 刘靖颖¹, 许雄彪^{1,2*}

(¹广西甘蔗生物学重点实验室, 广西大学, 南宁 530005; ²亚热带农业生物资源保护与利用国家重点实验室, 广西大学, 南宁 530004)

First report of a badnavirus infecting *Hibiscus tiliaceus* revealed by small RNA deep sequencing WANG Zhiyuan¹, FENG Ruying¹, LOU Yinian¹, LIANG Kaili¹, LIU Jingying¹, XU Xiongbiao^{1,2*} (¹Guangxi Key Laboratory for Sugarcane Biology, Guangxi University, Nanning 530005, China; ²State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, Guangxi University, Nanning 530004, China)

Abstract: Small RNA (sRNA) triggered RNA silencing plays significant roles in viral resistance response by plants. Analysis of virus-derived sRNA profiles in plants is an efficient method for virus identification and *de novo* assembly of viral genomes. In this study, suspected virus-infected *Hibiscus tiliaceus* samples collected from Nanning were used for sRNA library construction and subsequent deep sequencing. After the assembly of total sRNAs, *H. tiliaceus* leaves were found to be infected by hibiscus chlorotic mottle virus (HCMV). The library generated about 11.77 million sRNA reads, of which 15 093 can be mapped onto viral genomes. Using *de novo* assembly and GenBank Virus RefSeq database blast, a candidate virus covers 18.8% of the full length genome nucleotide sequence of a badnavirus, which has the highest similarity to hibiscus bacilliform virus isolate GD-1 (HBV-GD1). To confirm the existence of HCMV in the samples, a fragment of about 1 400 nucleotides encoding Reverse transcriptase /RNase H was obtained by PCR, and confirmed by Sanger sequencing. This is the first report about using sRNA deep sequencing technology to identify the *Hibiscus tiliaceus* infecting virus, and this provides important information for further study on the molecular characteristics of HCMV and virus-host interactions.

Keywords: small RNA; hibiscus chlorotic mottle virus; deep sequencing

中图分类号: S432.41 文献标识码: A 文章编号: 0412-0914(2024)01-0209-06

0 引言

黄槿 (*Hibiscus tiliaceus* L.) 为锦葵科 (*Malvaceae*) 木槿属 (*Hibiscus*) 常绿乔木或灌木, 广泛分布于台湾、福建、广西、广东等地, 是一种具有重要生态和观赏价值的半红树植物^[1]。同时黄槿具有清热解毒、散瘀消肿、止血镇痛抗炎等功效, 具有潜在的

药用价值^[2]。关于黄槿上的植物病害鲜有研究, 其上发生的病毒病害更是尚未见报道。RNA 沉默是植物抵御病毒入侵的重要途径^[3], 小 RNA 深度测序 (Small RNA Deep Sequencing) 作为高通量测序技术的重要应用, 可同时获得来源于病毒和宿主的小 RNA 序列, 是研究 RNA 沉默机制的有效工具。随

收稿日期: 2022-11-25; 修回日期: 2023-01-12 网络首发时间: 2023-03-30

网络首发地址: <https://kns.cnki.net/kcms/detail/11.2184.Q.20230329.1519.002.html>

基金项目: 广西大学高层次人才-助理教授科研启动经费项目 (A3310051012)

通信作者: 许雄彪, 博士, 讲师, 主要从事植物病理学研究; E-mail: xiongbiao@gxu.edu.cn

第一作者: 王致远, 硕士研究生, 研究方向资源利用与植物保护; E-mail: w1274105733@163.com;

冯如盈, 本科生, 研究方向植物保护; E-mail: 851021551@qq.com。

着小 RNA 深度测序深度的不断发展,小 RNA 深度测序数据高度重叠,能组装得到病毒全长基因组序列,其在病毒的鉴定和新病毒的发现方面具有广阔的前景,利用 sRNA 深度测序技术发现和鉴定作物和昆虫上的病毒已有广泛的应用^[4-5],现已成为一种快速鉴定植物病毒的有效手段。

本研究采用 sRNA 深度测序结合 PCR 技术,在广西南宁市三株表现褪绿白化、斑驳、皱缩的黄槿植株上鉴定到一种杆状 DNA 病毒,命名为黄槿褪绿斑驳病毒-南宁分离物 (hibiscus chlorotic mottle virus isolate Nanning, HCMV-NN),获得了该病毒的部分基因组序列,通过杆状 DNA 病毒属 RT/RNase H 保守序列相似性分析及构建系统进化树,明确了 HYMV-NN 与其它杆状 DNA 病毒之间的进化关系。同时,利用生物信息学的方法对该病毒来源的 sRNA 进行了特征分析。

1 材料与方法

1.1 实验材料

2020 年 12 月,从广西壮族自治区南宁市青秀区采集 3 份疑似病毒侵染的黄槿叶片样品,感病植株叶片表现褪绿、白化、斑驳、卷曲、坏死等症状 (图 1)。样品采集后迅速置于液氮中冻存,置于 -80 ℃ 冰箱保存。

1.2 植物总 RNA 提取及小 RNA 深度测序分析

采用 Trizol 法提取黄槿叶片的总 RNA,经 Nanodrop-2000 (Thermo Scientific) 检测 RNA 纯度。将质量合格的 RNA 样品送至北京诺禾致源科技股份有限公司 Illumina HiSeq 2000 平台进行深度测

序。运用生物信息学手段对原始数据进行过滤处理:去除接头序列,去除无法确定碱基信息比例大于 10% 的序列,去除低质量碱基,去除未插入 3' 接头、5' 接头的 reads,获得不含接头序列的 sRNA 序列,在此基础上筛选出 18~26 nt 的 sRNA 序列。

1.3 测序数据的序列拼接、BLAST 分析及病毒相关序列的筛查

分别采用 PFOR 软件和 velvet 软件对过滤后的 clean reads 进行 sRNA 拼接^[6-7],再将两者重叠部分进行 merge,以获得更准确的 contigs。得到的 contigs 用 Basic Local Alignment Search Tool (BLAST) 进行比对并注释,经过 BLAST 同源比对筛查到与之相似性最高的病毒全长基因组序列,并以此作为参考基因组,应用生物信息手段比对得到病毒来源的 sRNA,并用 Geneious Prime 软件对 sRNA 映射在病毒基因组上的位置进行分析。

1.4 PCR 扩增、克隆及序列分析

利用改良 CTAB 法提取黄槿叶片总 DNA,以此为模板,根据拼接到的病毒基因组序列设计特异性引物 HCMV ORF3 Partial/F:5'-CTGAAGTATG-GCCAGATG-3',HCMV ORF3 Partial/R:5'-ACTC-CTCCCCATCCTTCCATAC-3',用 Phanta EVO HS Super-Fidelity DNA Polymerase (Vazyme) 进行 PCR 扩增,PCR 产物纯化后连接至 pCE2 TA/Blunt-Zero Vector (Vazyme) 载体,并转化 *Escherichia coli* 菌株 DH5α 感受态细胞,涂布于含氨苄青霉素的 LB 固体培养基上,培养过夜后挑选单菌落,用 2 × Phanta Max Master Mix (Vazyme) 扩增鉴定阳性克隆



Fig. 1 *Hibiscus tiliaceus* leaves showing bleaching, mottle, blade curling and necrosis symptoms

随机挑取 2 个阳性克隆送至上海生工生物工程股份有限公司进行测序,获得的序列利用 BLAST 进行比较分析。使用 MEGA7.0 进行多重序列相似性分析,并用邻接法(Neighbor-Joining, NJ)构建系统进化树。

2 结果

2.1 感病黄槿样品小 RNA 深度测序结果

该感病黄槿样品深度测序后得到的 raw reads 经过过滤后共得到 14 898 564 条 clean reads,对长度为 18~26 nt 的有效 reads 进行筛选,得到的 contigs 为 11 770 925 条,占总 sRNA Reads 数量的 79.01%。用 velvet 软件拼接组装(K-mer 值为 17) 共获得长度<50 nt 的 contigs 168 条,长度介于 50~100 nt 的 contigs 1 631 条,长度介于 100~200 nt 的 contigs 84 条,长度介于 200~500 nt 的 contigs 5 条。用得到的 contigs 在 NCBI 中进行比对(BLASTn),发现共 27 条 contigs 能比对到病毒基因组,其中有 2 条 contigs 能比对到葡萄叶褪色相关病毒(grapevine Roditis leaf discoloration-associated virus, GRLDaV, GenBank 登录号: NC_027131.1),有 10 条 contigs 能比对到朱槿杆状病毒广东分离物 1(hibiscus bacilliform virus, HBV-GD1, GenBank 登录号: KF875586.1),具有最高的丰度(37.0%),拼接得到的序列总长度覆盖 HBV-GD1 基因组全长的 18.8%(图 2)。

2.2 HCMV-NN 来源的 sRNA 特征分析

将上述 11 770 925 条长度介于 18~26 nt 的 sRNA 进行 bowtie 过滤分析,允许 1 个错配的 sRNA Reads 为 1 885 161,与 GenBank Virus RefSeq 核酸数据库进行比对,发现 15 093 条 sRNA 能匹配到病毒基因组。病毒侵染的植物, RNA 沉默产生的 viral siRNA(vsiRNA) 长度主要为 21、22 和 24 nt 三个类群^[8]。对不同长度的 vsiRNA 进行统计分析,发现长度为 21、22 和 24 nt 的 sRNA 具有较高的丰度,占比分别为 24.78、17.81 和 28.88%,三者总占比为 71.47%(图 3)。

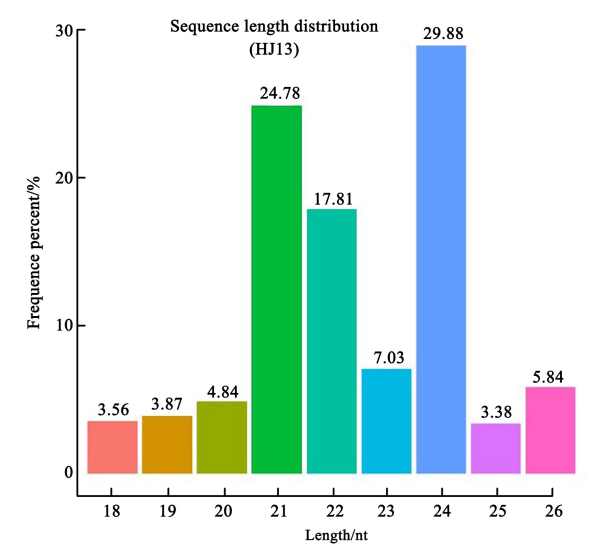


Fig. 3 Size distribution of HCMV-NN derived sRNAs



Fig. 2 Size distribution of Contigs that mapped on the reference viral genome

2.3 HCMV-NN 基因组部分序列的测定与分析

将 PFOR 和 velvet 拼接结果重叠部分进行 merge 分析,以 HBV-GD1 基因组序列作为参考基因组,将拼接得到的 contigs map 到 HBV-GD1 参考基因组上,根据拼接的 contigs 序列设计一对特异性引物进行 HCMV 部分基因组的扩增,可扩增到长度为 1 397 bp 的病毒特异性序列(图 4)。经测序及同源比对分析,发现该病毒特异性扩增序列与 HBV-GD1 病毒序列相似性为 76.3%,这段序列位于 HBV-GD1 病毒 ORF3 编码的 Reverse transcriptase /RNase H (RT/RNase H) 区域。根据国际病毒分类委员会(International Committee on Taxonomy of Viruses: ICTV)将 *Badnavirus* 病毒 RT/RNase H 序列差异大于 20%作为种之间的分类标准^[9-10],该病毒分离物与已知的 HBV-GD1 的 RT/RNase H 序列差异为 27.2%,说明该病毒与 *Hibiscus bacilliform virus* 为两个不同种,根据其在寄主植物黄槿上引起的症状,暂命名为黄槿褪绿斑驳病毒-南宁分离物(*hibiscus chlorotic mottle virus isolate Nanning*, HCMV-NN)。

2.4 HCMV-NN 的系统进化分析

因 HCMV-NN 属于杆状 DNA 病毒,且与 HBV-GD1 具有较高的序列相似性,用邻接法(Neighbor-Joining method)对包括 HBV-GD1 在内的已知的 19 条 *Badnavirus* 基因组(Table 1) RT/

RNase H 区域对应位置长度约 1 400 bp 序列进行系统发育分析(图 5),系统进化树显示 HCMV-NN 与 HBV-GD1 亲缘关系最近,且与柑橘黄化花叶病毒(*Citrus yellow mosaic virus*, CMBV, GenBank 登录号: NC_003382)形成一个小分支,而与其他 *badnavirus* 亲缘关系较远。

3 讨论

由于木本植物病毒滴度较低,难以提纯获得完整的病毒粒子,使得依赖生物学特性、基因组信息、血清学特性、理化特性等信息的传统病毒学检测方法来检测未知木本植物病毒受到极大限制。小 RNA 深度测序作为高通量测序技术的重要应用,可获得病毒来源的 sRNA (vsiRNA) 序列,并能组装得到病毒全基因组序列^[4, 8],在病毒的鉴定和新病毒的发现方面应用越来越广泛。但由于 vsiRNAs 较短且丰度存在较大差异,通过 *De novo* 组装获取完整病毒基因组序列仍存在一定挑战。同时,由于病毒变异快和高通量测序的误差等原因,对 sRNA 深度测序组装得到的病毒序列还需进一步实验验证。

本文利用小 RNA 深度测序和组装技术,将分离自南宁的黄槿褪绿斑驳病的病原鉴定为 HCMV-NN,并获得了其 ORF3 中长度为 1 397 bp 的 RT/RNase H 部分保守序列。发现感染 HCMV 的黄槿在抵抗病毒侵染过程中产生的 siRNA 长度主要

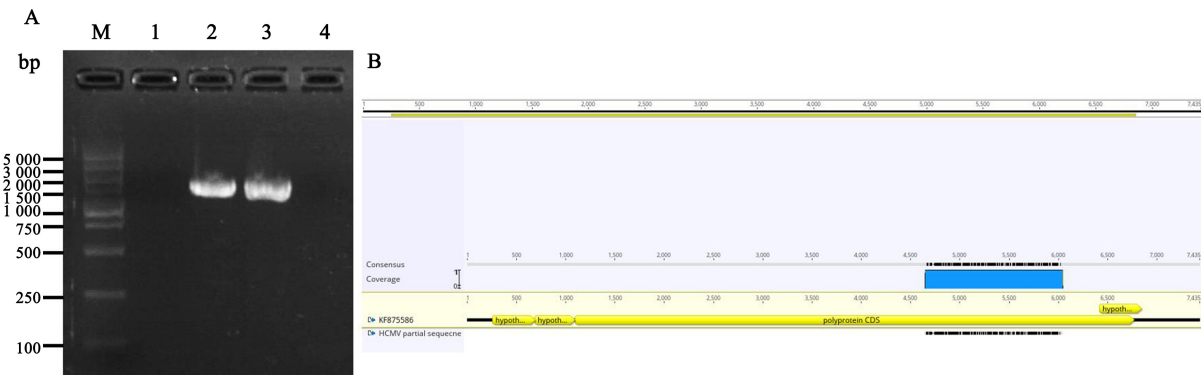


Fig. 4 The amplified products of HCMV-NN and their distribution on the reference viral genome

A; Lane 1: Negative control using healthy leaves; 2, 3: Partial sequences of HCMV-NN amplified using prime pair HCMV-ORF3 partial-F/R; 4: Mock control; M: DNA marker DL5000. B: Distribution of the PCR products on HBV-GD1 reference viral genome.

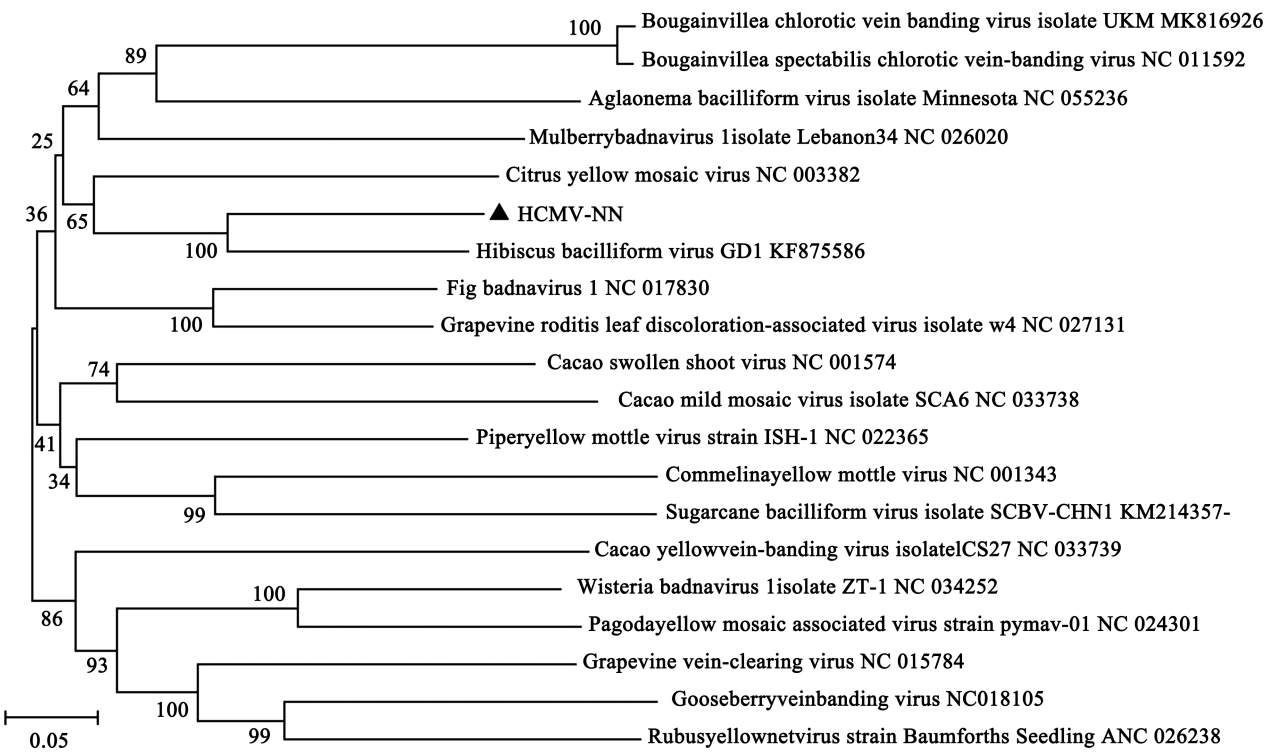


Fig. 5 Phylogenetic tree based on RT/RNase H nucleotides of hibiscus chlorotic mottle virus and known badnaviruses using neighbor-joining method

▲: Hibiscus chlorotic mottle virus isolate Nanning.

Table 1 The conserved RT/RNase H region from other badnaviruses

| Source | Location on genome | GenBank accession number |
|--|--------------------|--------------------------|
| Aglaonema bacilliform virus isolate Minnesota | 4 355-5 736 | NC_055236 |
| Bougainvillea chlorotic vein banding virus isolate UKM | 6 006-7 418 | MK816926 |
| Bougainvillea spectabilis chlorotic vein-banding virus | 5 949-6 089 | NC_011592 |
| Commelina yellow mottle virus | 5 283-6 676 | NC_001343 |
| Cacao swollen shoot virus | 4 623-6079 | NC_001574 |
| Cacao mild mosaic virus isolate SCA6 | 4 845-6 244 | NC_033738 |
| Cacao yellow vein-banding virus isolate ICS27 | 4 948-6 346 | NC_033739 |
| Citrus yellow mosaic virus | 4 816-6 621 | NC_003382 |
| Fig badnavirus 1 | 4 436-5 526 | NC_017830 |
| Grapevine vein-clearing virus | 5 350-6 731 | NC_015784 |
| Grapevine roditis leaf discoloration-associated virus isolate w4 | 4 373-5 763 | NC_027131 |
| Gooseberry vein banding virus | 5 293-6 677 | NC_018105 |
| Hibiscus bacilliform virus GD1 | 4 644-6 043 | KF875586 |
| Mulberry badnavirus 1 isolate Lebanon34 | 4 259-5 646 | NC_026020 |
| Piper yellow mottle virus strain ISH-1 | 5 275-6 662 | NC_022365 |
| Pagoda yellow mosaic associated virus strain pymav-01 | 4 742-6 167 | NC_024301 |
| Rubus yellow net virus strain Baumforth's Seedling A | 5 202-6 586 | NC_026238 |
| Sugarcane bacilliform virus isolate SCBV-CHN1 | 4 835-6 266 | KM214357 |
| Wisteria badnavirus 1 isolate ZT-1 | 4 639-6 065 | NC_034252 |

为 21、22 和 24 nt, 其中 24 nt 的 siRNA 丰度最高 (28.88%), 推测 DCL4、DCL2 和 DCL3 可能参与了黄槿抗病毒 RNA 沉默过程, 且 DCL3 在此过程中起主要的作用。

sRNA 拼接的得到的 contigs 均 mapped 到参考基因组 ORF3 上, 特别在 RT/RNase H 编码区域, 说明 ORF3 可能为 HCMV vsiRNA 产生的热点区域, 且黄槿对 HCMV RT/RNase H 编码区域具有强烈的剪切作用。杆状 DNA 病毒 ORF3 通过编码一个多聚蛋白切割产生 Reverse transcriptase 和 RNase H 完成自身病毒基因组的复制, 而 vsiRNA 主要来源于该区域, 推测黄槿主要通过抑制 HCMV 的复制而抵御病毒的侵染, 拓宽了对杆状病毒属病毒-寄主相互作用的认知。关于 HCMV 病毒在抵抗寄主 RNA 沉默过程中编码的 RNA 沉默抑制子及具体作用机制仍需进一步研究。

参考文献

- [1] ZHANG W W, LIU N, WANG J, *et al.* Ecological and biological characteristics of *Hibiscus tiliaceus*, a mangrove associate in China (in Chinese) [J]. *Guihaia* (广西植物), 2012, 32(2): 5.
- [2] SHAO C L, FU X M, WANG C Y, *et al.* Investigation on the Status of Mangrove Resources and Medicinal Research in China III. Status of Folk Medicinal Usage and Medicinal Research (in Chinese) [J]. *Periodical of Ocean University of China* (中国海洋大学学报), 2009, 39(4): 712-718.
- [3] ZHAO Y L, YANG X, ZHOU G H, *et al.* Engineering plant virus resistance: From RNA silencing to genome editing strategies [J]. *Plant Biotechnology Journal*, 2020, 18(2): 328-336.
- [4] WU Q F, LUO Y J, LU R, *et al.* Virus discovery by deep sequencing and assembly of virus-derived small silencing RNAs [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2010, 107(4): 1606-1611.
- [5] XU Y, HUANG L Z, WANG Z C, *et al.* Identification of Himetobi P virus in the small brown planthopper by deep sequencing and assembly of virus-derived small interfering RNAs [J]. *Virus Research*, 2014, 179: 235-240.
- [6] ZERBINO D R, BIRNEY E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs [J]. *Genome Research*, 2008, 18: 821-829.
- [7] AREGGER M, BORAH B K, SEGUIN J, *et al.* Primary and secondary siRNAs in geminivirus-induced gene silencing [J]. *PLoS Pathogens*, 2012, 8(9): e1002941.
- [8] NIU X R, SUN Y, CHEN Z, *et al.* Using small RNA-seq data to detect siRNA duplexes induced by plant viruses [J]. *Genes*, 2017, 8(6): 163.
- [9] FAUQUET C M, MAYO M A, MANILOFF J, *et al.* Virus taxonomy, 7th report of the ICTV [C]. London: Elsevier, 2005.
- [10] KING A M Q, ADAMS M J, CARSTENS E B, *et al.* Virus taxonomy, 9th report of the ICTV [C]. London: Elsevier, 2011.

责任编辑: 张宗英