

草莓白化相关病毒中国分离物全基因组分析

陈 道, 张 洁, 吴祖建, 丁新伦*

(福建农林大学植物病毒研究所, 闽台作物有害生物生态防控国家重点实验室, 福建省植物病毒重点实验室, 福州 350002)

摘 要: 草莓白化相关病毒 (strawberry pallidosis-associated virus, SPaV) 属于长线形病毒科 (Closteroviridae) 毛形病毒属 (Crinivirus), 可引起草莓病害, 2017 年在中国首次报道。采用高通量测序、RACE 和 RT-PCR 技术获得了 SPaV 中国分离物 (FJ) 的基因组全长。该病毒含有两条正单链基因组 RNA1 和 RNA2。RNA1 全长 8 048 nt, 5'和 3'非编码区序列分别为 264 和 197 nt, 含有 3 个开放阅读框(ORF), 分别编码 ORF 1a/1b 融合蛋白和 p9 蛋白。RNA2 全长 7 977 nt, 5'和 3'非编码区序列分别为 248 和 186 nt, 含有 8 个开放阅读框 (ORF), 分别编码 HSP70h、CPh、CP、CPm、p7、p6、p9 和 p28 等 8 个蛋白。RNA1 和 RNA2 与美国 M1 分离物分别具有 98.5%和 99.0%的核苷酸一致性; 系统发育分析结果表明, SPaV 中国分离物 (FJ) 单独处在一个分支。对 SPaV 来源的小 RNA 的分析表明, 来源于 SPaV 的小 RNA 长度以 21 和 22 nt 为主。

关键词: 草莓; 草莓白化相关病毒; 高通量测序; RT-PCR; 基因组序列

中图分类号: S 668.4

文献标志码: A

文章编号: 0513-353X (2021) 01-0146-07

Genome Sequence of Strawberry Pallidosis-associated Virus Isolated in China

CHEN Dao, ZHANG Jie, WU Zujian, and DING Xinlun*

(State Key Laboratory of Ecological Pest Control for Fujian and Taiwan Crops, Key Laboratory of Plant Virology of Fujian Province, Institute of Plant Virology, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract: Strawberry pallidosis-associated virus (SpaV) belongs to the genus *Crinivirus* of the family *Closteroviridae*, which causes strawberry diseases. It was first reported in China in 2017. In this study, high-throughput sequencing, RACE and RT-PCR techniques were used to obtain the full length of SPaV China isolate (FJ). The virus contains two positive-sense single, stranded genome RNA1 and RNA2, RNA1 is composed of 8 048 nt and contains three open reading frames (ORFs), encoding ORF 1a/1b fusion protein and p9 protein, flanked by 5' and 3' untranslated regions of 264 nt and 197 nt, respectively. RNA2 is composed of 7 977 nt and contains eight ORFs, encoding HSP70h, CPh, CP, CPm and other four unknown function proteins (p7, p6, p9 and p28), flanked by 5' and 3' untranslated regions of 248 and 186 nt, respectively. RNA1 and RNA2 share 98.5% and 99.0% nucleotide sequence identity with that of M1 isolate (USA), respectively. Phylogenetic analysis showed that SPaV China isolate (FJ) was clustered

收稿日期: 2020-09-13; 修回日期: 2020-12-11

基金项目: 福建省自然科学基金项目 (2019J01656); 福建农林大学科技创新专项 (CXZX2016132, CXZX2019018G)

* 通信作者 Author for correspondence (E-mail: dingxinlun@163.com)

in a separate branch. The analysis of sRNA derived from SPaV showed that the dominant sizes were 21 nt and 22 nt.

Keywords: *Fragaria* × *ananassa*; strawberry pallidosis-associated virus; NGS; RT-PCR; viral genome sequence

中国目前草莓的栽培面积和产量均居世界首位。据统计, 2018 年中国草莓的播种面积为 119.97 千 hm^2 , 产量达到 306.03 万 t (国家统计局数据: <http://data.stats.gov.cn>)。生产上采用匍匐茎繁殖、种苗地区间的调运以及种植面积的迅速扩大等使得草莓病毒病发生面积上升, 病毒病害已成为制约中国草莓生产的重要因素。已报道约 30 种病毒可以引起草莓病害, 其中草莓镶脉病毒 (strawberry vein banding virus, SVBV)、草莓皱缩病毒 (strawberry crinkle virus, SCV)、草莓斑驳病毒 (strawberry mottle virus, SMoV) 和草莓轻型黄边病毒 (strawberry mild yellow edge virus, SMYEV) 等 4 种由蚜虫传播的草莓病毒在世界各地广泛分布, 危害严重。近年来, SCV 的发生部分地区有下降的趋势。王运生等 (2015) 对北京、浙江、河北和四川等地的调查结果表明 SCV 检出率较低; 福建和新疆等地未检测到 SCV (陈道 等, 2018; 杨波 等, 2018)。近年来中国也相继报道了一些草莓新病毒, 如草莓坏死休克病毒 strawberry necrotic shock virus, SNSV (Li & Yang, 2011)、黄瓜花叶病毒 cucumber mosaic virus, CMV (Chen et al., 2014)、草莓白化相关病毒 strawberry pallidosis-associated virus, SPaV (Ding et al., 2017; Shi et al., 2018)、草莓毛形病毒属病毒 3 号 strawberry crinivirus 3 (Chen et al., 2018)、草莓毛形病毒属病毒 4 号 strawberry crinivirus 4 (Chen et al., 2018) 以及最近发现的草莓病毒——草莓相关病毒 1 号 strawberry-associated virus 1 (Ding et al., 2019)。由于大多数草莓病毒病由介体昆虫传播, 生产上目前仍采用“治虫防病”的策略来防控草莓病毒病害的发生。

草莓白化相关病毒 (SPaV) 属于长线形病毒科 (*Closteroviridae*) 毛形病毒属 (*Crinivirus*), 可引起白化相关病毒病害 (Pallidosis disease), 最早于 1950 年代在美国和澳大利亚报道, 中国 2017 年在福建首次报道 (Ding et al., 2017), 2018 年在河南被报道 (Shi et al., 2018)。SPaV 可通过嫁接和介体昆虫温室白粉虱 (*Trialeurodes vaporariorum*) 传播。弗吉尼亚草莓 (*F. virginiana*) 两个单系 (UC-10 和 UC-11) 受侵染后表现叶片畸形、黄化和植株矮化等症状, 而森林草莓 (*F. vesca*) 不表现症状。虽然单一病毒侵染商品化草莓几乎不表现明显症状, 但其与其他草莓病毒的复合侵染通常会引起草莓发生严重病害 (Martin & Tzanetakis, 2006)。该病毒基因组含有 2 条正单链 RNA (RNA1 和 RNA2), RNA1 全长 8 066 nt, 至少包含 3 个开放阅读框 (open reading frame, ORF), 其中 ORF1a/1b 为多功能蛋白, 含有木瓜样蛋白酶、甲基转移酶和 RNA 解螺旋酶结构域, ORF1b 编码 RNA 依赖的 RNA 聚合酶 (RdRp), 3' ORF 编码含有 2 个跨膜螺旋的小蛋白; RNA2 全长 7 978 nt, 包含 8 个 ORF, 推测编码运动蛋白、结构蛋白以及其他未知功能蛋白 (Tzanetakis et al., 2005)。SPaV 各编码蛋白的功能还需进一步验证。

2017 和 2018 年中国福建和河南分别率先报道发现 SPaV, 该病毒在两省的发病率均较低 (Ding et al., 2017; Shi et al., 2018), SPaV 在中国其他地区尚未见报道, 该病毒在中国的分布、发病情况和对草莓产业的危害还需调研和评估。此外, SPaV 和其他病毒复合侵染时可引起严重的病毒病害, 中国 SPaV 和其他病毒复合侵染的情况以及复合侵染的致病机制也值得进一步研究。目前对于 SPaV 中国分离物的基因组信息尚不清楚。本研究中采用高通量测序技术 (Next-generation sequencing, NGS)、RACE 和 RT-PCR 的方法获得其基因组序列全长并对其进行分析, 以期 SPaV 蛋白功能、致病机理以及抗病品种筛选等研究奠定基础。

1 材料与amp;方法

1.1 供试材料及其小 RNA 和总 RNA 提取

草莓病株于 2017 年采自福建省福州市商品化草莓种植地。病叶于 - 70 ℃ 保存。称取 0.1 g 草莓病叶，利用 mirPremier microRNA Isolation Kit (Sigma-Aldrich) 试剂盒提取小 RNA；称取 0.1 g 草莓病叶，利用 RNAprep Pure 多糖多酚植物总 RNA 提取试剂盒提取总 RNA。

1.2 小 RNA 文库构建和高通量测序

采用 NEBNext Multiplex Small RNA library Prep Set for Illumina (NEB) 试剂盒构建 small RNA 文库。使用 Illumina HiSeq 平台测序，测序长度 1 × 50 bp，由苏州金唯智生物科技有限公司完成。对测序原始数据通过 Trimmomatic (v0.30) 去除接头以及低质量序列。使用短序列比对软件将序列比对到宿主的参考基因组上，并统计比对结果，保留没有比对到宿主上的序列，以排除宿主 RNA 序列。使用去除宿主 sRNA 后的 read，利用软件 velvet (EMBL-EBI, UK) 结合 SeqMan software (Lasergene) 进行组装，各样本拼接结果分别进行病毒蛋白数据库和病毒参考基因组数据库 Blast 注释。为验证注释到的病毒的准确性，选取所有注释到的病毒作为候选病毒，下载所有候选病毒的参考基因组，利用短序列比对软件将待分析数据比对到候选病毒的参考基因组上，统计每个样本在每个候选病毒基因组上的覆盖度及覆盖深度。

1.3 RT-PCR 扩增

采用 RT reagents 反转录试剂盒 (TaKaRa) 合成单链 cDNA，反应条件：42 ℃ 15 min，85 ℃ 5 s。以适量的 cDNA 为模板，PCR 扩增病毒各基因组片段，循环参数为：95 ℃ 10 s，95 ℃ 10 s，47 ~ 67 ℃ 退火 10 s，36 个循环。引物依据 NGS 获得的病毒序列设计，引物序列见表 1。PCR 产物经电泳回收、连接至 pEASY-Blunt Simple 克隆载体 (北京全式金生物技术有限公司)，并转化 DH5α 大肠杆菌，阳性克隆子经 PCR 鉴定后由铂尚生物技术有限公司测序。采用 SMARTer RACE 5'/3'试剂盒 (Clontech) 分别扩增病毒的 5'和 3'末端序列。

表 1 本研究中使用的引物
Table 1 Primers used in this study

引物名称 Primer name	引物序列 (5' - 3') Primer pair	退火温度/℃ Temperature
RNA1-1	F: GTAATGAAGTCGTCGAGACCA; R: GAATAGACTACTTCGCCATGTCTG	47
RNA1-2	F: GAGACGATAACGATGTGCGAC; R: TCGACAGTTGAGAGTCGTGAC	52
RNA1-3	F: TCTAAATTTAAGAGTTTGCTGTTTAA; R: ATAGCTCTAGAGGCTCTAGCAIT	57
RNA1-4	F: TCGGTATCAGTATGTCACCAC; R: CATGGCCAAGGTCACCAGAGT	56
RNA1-5	F: TCTTATAATTCATACCCTTGAAC; R: CGGGAATTTATTGTAAATCATAC	58
race1-3	ACACTGGAGTTAAAGTCTGAAACTCTAAATTTAAGAGTTTAG	57
race1-5	TTTAAAAGTAACTGATACATCTTCAGGCAGACATAATCTCT	61
RNA2-1	F: CCAAAGATAAAATAACAAACATCAA; R: ATATCTTCCGCATACGAATTTATCAA	57
RNA2-2	F: GGGCAGATCAAAGAACAACCTTGTT; R: CATTGATCGTTCCACTCAATTTGAAA	60
RNA2-3	F: ATGACGTCGCAAATGATAGTGCTTTC; R: CTTCCAACCTCTGCTCGGTAACCT	55
RNA2-4	F: CTTGATAAACGTGACAAATCAAACCT; R: GGGGTCAGCTCTTTCAAAGCTTGA	57
RNA2-5	F: CTTTTTATTAGCGTCAATCTTAGCTTT; R: TTTCTCGGTCAATCGATCGAACTCTGA	58
RNA2-6	F: CAAGAATAATGGATTGGAGGATATGCCTT; R: TCACCAGGGGATTGCTTTCCAACCT	56
race2-3	GATTACGCCAAGCTTTGCTACCTCCTATACCAATGCGGA	67
race2-5	ACCTCCCAAACGAATTTATCAATCTCATAAGGTAAAGTGTT	65

1.4 SPaV 基因组序列分析

采用 ORF Finder 在线软件对获得的 SPaV 基因组结构进行分析。采用 Mega 6.0 软件, 基于 Tamura-Nei 模型的最大似然法 (Maximum Likelihood, ML) 构建基于 SPaV CP 基因的系统进化树, 以自展法 (Bootstrap) 检验系统树, 自展重复次数设定为 1 000 次, 以莴苣侵染性黄化病毒 (lettuce infectious yellows virus, LIYV) 的 CP 基因作为外类群。

2 结果与分析

2.1 感病草莓的小 RNA 高通量测序

草莓病叶经 NGS 测序获得原始 read 为 1 649 826 744 条, 样品测序数据过滤后获得 24 487 004 条 read。对获得序列的长度和数量进行分析, 小 RNA 的长度主要分布在 18 ~ 32 nt, 其中长度为 20 ~ 24 nt 的小 RNA 含量较高。使用去除宿主 sRNA 后的 read 进行组装, SPaV RNA1 覆盖长度为 8 016 nt, 覆盖率为 99.38%, 测序深度为 249.12; RNA2 覆盖长度为 7 978 nt, 覆盖率为 99.65%, 测序深度为 936.72。

2.2 SPaV 基因组全长序列分析

为了确认 NGS 的结果, 并补足测序无法覆盖到的病毒区域和两端序列, 以 NGS 测序获得的部分基因组序列设计引物, RT-PCR 扩增得到长度为 500 ~ 2 000 bp 的 SPaV 序列片段, 5' RACE 和 3' RACE 得到 SPaV 的 5'和 3'末端序列, 各基因组片段见图 1。各片段组装后获得 SPaV 的基因组 RNA1 和 RNA2; 将 RNA1 和 RNA2 基因组序列提交至 GenBank 数据库, 其登录号分别为 MN747001 和 MN747002。

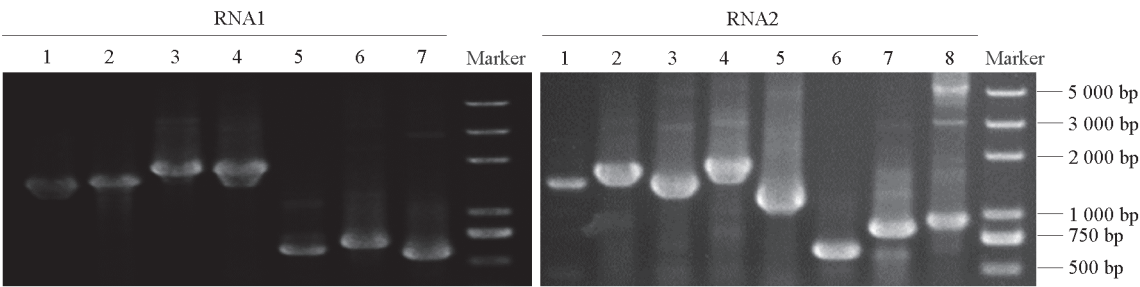


图 1 SPaV 的 RNA1 和 RNA2 基因组片段
Fig. 1 RNA1 and RNA2 fragments of SPaV

RNA1 全长 8 048 nt, 与美国 M1 分离物 (NC_005895) 核苷酸序列同源性为 98.5%, 5'和 3'非编码区序列分别为 264 和 197 nt, RNA1 基因组结构如图 2 所示, 含有 3 个 ORF, 其中 ORF1 和 ORF2 采用 +1 核糖体移码策略, 编码 ORF 1a/1b 融合蛋白 (nt 265 ~ 6 082、6 084 ~ 6 764、6 768 ~ 7 603), ORF1a 含有木瓜蛋白酶 (P-Pro)、甲基转移酶 (Mtr) 和解螺旋酶 (Hel) 结构域, ORF1b 编码 RNA 依赖 RNA 聚合酶 (RdRP), 而 ORF3 (nt 7 612 ~ 7 851) 编码大小为 9 kD 蛋白, 功能未知。在基因组 nt 3 983 ~ 3 984 区间, 缺失 18 nt 核苷酸序列 (CTTCCTGAACGTTGGCTG), 该缺失区域位于 ORF 1a/1b 融合蛋白基因内, 少编码 6 个氨基酸 (FLNVGC)。

RNA2 全长 7 977 nt，与美国 M1 分离物（NC_005896）有 99%的核苷酸同源性，5'和 3'非编码区序列分别为 248 和 186 nt。其中 5'末端具有核苷酸序列 GGTAAT，而美国分离物则不含有该序列。RNA2 基因组结构如图 2 所示，含有 8 个 ORF，分别编码类热激蛋白（HSP70h, nt 677 ~ 2 347）、类衣壳蛋白（CPh, nt 2 509 ~ 4 065）、衣壳蛋白（CP, nt 4 294 ~ 5 040）、次衣壳蛋白（CPm, nt 5 009 ~ 7 078）以及其他 4 个功能未知蛋白 p7（nt 249 ~ 437）、p6（nt 2 354 ~ 2 515）、p9（nt 4 044 ~ 4 301）和 p28（nt 7 081 ~ 7 791）。在基因组 nt 7 409 ~ 7 410 区间缺失 9 nt 核苷酸序列（TTGATGTTA），该缺失区域位于 p28 内，少编码 3 个氨基酸（LML）。

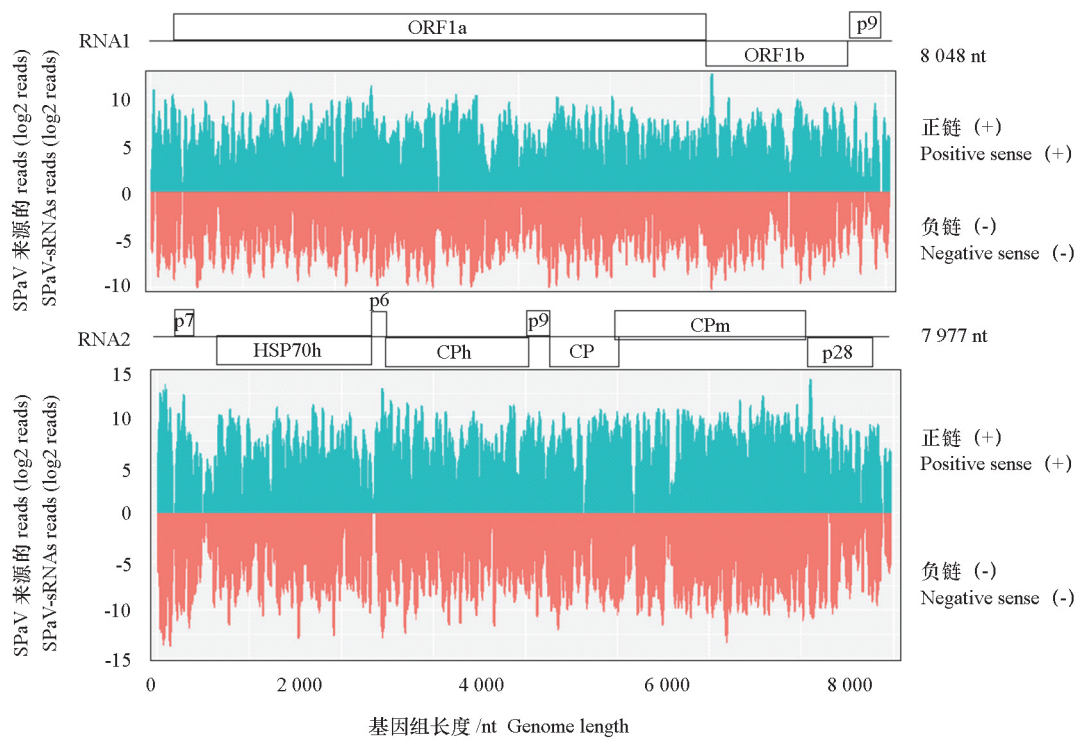


图 2 SPaV 基因组结构以及小 RNA 在 SPaV 基因组上的分布
Fig. 2 Genome organization of the SPaV and sRNA distribution

2.3 SPaV 中国分离物与其他分离物亲缘关系分析

以 SPaV 中国分离物（福建，FJ）和其他分离物的外壳蛋白（CP）核苷酸序列构建了系统发育树，以毛形病毒属的模式种 LIYV 作为外群，结果（图 3）表明，6 个 SPaV 分离物可分为两个大的进化分支，其中 SPaV 中国河南分离物（ZM4 和 ZM44）和美国分离物（CFRA 9064、C1 和 M1）归为一个大的分支，而 SPaV 中国福建分离物（FJ）单独处在一个分支。

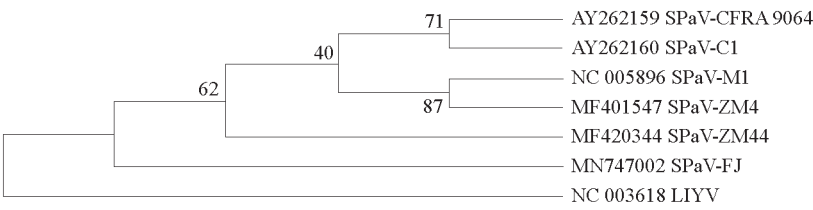


图 3 基于不同来源 SPaV 分离物 CP 核苷酸序列构建的系统进化树
Fig. 3 Maximum-likelihood phylogenetic tree based on the CP nucleotide sequences of SPaV

2.4 来源于 SPaV 的 sRNA 分析

共有 563 793 read 匹配到 SPaV 基因组序列, 其中匹配到 RNA1 和 RNA2 的 read 分别为 118 215 和 445 578。来源于 SPaV 的 sRNA 在基因组上的分布如图 2 所示, sRNA 在 RNA1 和 RNA2 整个基因组序列区域均有分布, 而且分布比较均匀, 匹配到基因组 RNA1 正链和负链的 sRNA 分别为 48 925 和 69 290; 匹配到基因组 RNA2 正链和负链的 sRNA 分别为 222 316 和 223 262; 对 SPaV 来源的小 RNA 进行长度分析, 18 ~ 26 nt sRNA 的数量分别是 15 422、34 276、74 209、197 226、194 422、32 570、8 895、2 146 和 1 302 条, 不同长度 sRNA 所占比例见图 4, 其中 21 和 22 nt 的 sRNA 占主要部分, 所占比例均为 34%。

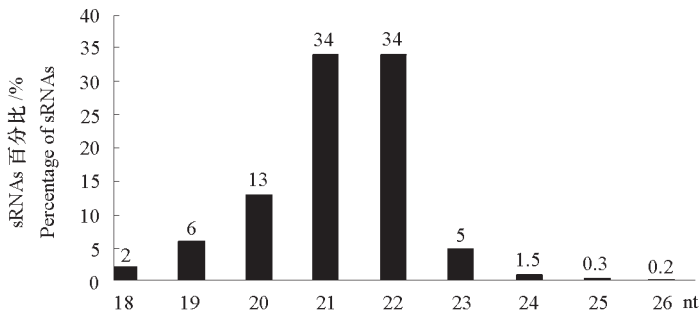


图 4 SPaV 来源的小 RNA 的长度分布
Fig. 4 Size distribution of SPaV derived sRNAs

3 讨论

采用 NGS 和 RT-PCR 结合获得了 SPaV 中国分离物 (FJ) 的基因组全序列, 序列分析表明 SPaV 中国分离物和美国分离物基因组具有较高的核苷酸序列一致性, 两者的基因组结构相似。值得注意的是, SPaV 中国分离物 (FJ) 基因组的 RNA2 单链 5' 末端含有 8 个核苷酸序列 GGTA AAAAT, 美国 M1 分离物则不具有此序列, 该序列的存在可能对病毒的复制和转录有影响; 此外, 位于 RNA1 单链 ORF 1a/1b 融合蛋白和位于 RNA2 单链 p28 基因编码区均缺失了核苷酸序列, 导致病毒蛋白少编码了氨基酸, 核苷酸的缺失是否会影响编码蛋白功能以及病毒的致病性还有待研究。SPaV 全基因组序列的获得可为病毒分子检测技术的研发、病毒蛋白功能以及致病机理等方面的研究奠定基础。亲缘关系分析表明, SPaV 河南分离物和美国分离物归为一个大的分支, 二者亲缘关系较近, 福建

分离物单独形成一个分支,表明福建分离物同河南分离物以及美国分离物亲缘关系较远,该福建分离物和河南分离物的病毒毒源可能来自于不同的地区。

对 SPaV 来源的 sRNA 的分析结果显示, sRNA 在基因组上呈均匀分布;来自病毒 RNA1 的正链和负链的 sRNA 数量相当,推测这些 sRNA 可能来自于病毒复制中间体。SPaV 来源的 sRNA 长度以 21 和 22 nt 为主,推测两种核酸内切酶 DCL4 和 DCL2 在草莓的抗病毒中发挥了主要作用,这为研究病毒来源的小 RNA 在抗病毒机制中的功能以及寻求草莓病毒病的防治策略提供理论基础。

References

- Chen D, Ding X L, Wang A M, Zhang J, Wu Z J. 2018. First report of strawberry crinivirus 3 and strawberry crinivirus 4 on strawberry in China. *New Disease Reports*, 37: 24.
- Chen Dao, Ding Xing-lun, Zhang Jie, Xie Li-yan, Wu Zu-jian. 2018. Molecular detection of six viruses infecting strawberry in Fujian Province. *Journal of Plant Protection*, 45 (6): 1433 - 1434. (in Chinese)
- 陈 道, 丁新伦, 张 洁, 谢荔岩, 吴祖建. 2018. 福建省六种草莓病毒的分子检测. *植物保护学报*, 45 (6): 1433 - 1434.
- Chen L, Shang Q X, Chen X Y, Xing D M, Yang R, Han C G, Ran C, Wei Y M, Zhao X Y, Liu Z P. 2014. First report on the occurrence of *Cucumber mosaic virus* on *Fragaria ananassa* in China. *Plant Disease*, 98 (7): 1015.
- Ding X L, Chen D, Du Z G, Zhang J, Wu Z J. 2019. The complete genome sequence of a novel cytorhabdovirus identified in strawberry (*Fragaria ananassa* Duch.). *Archives of Virology*, 164 (12): 3127 - 3131.
- Ding X L, Chen D, Wang A M, Zhang J, Wu Z J. 2017. First report of *strawberry pallidosis-associated virus* on strawberry in China. *Plant Disease*, 101 (12): 2154.
- Li L, Yang H. 2011. First report of *Strawberry necrotic shock virus* in China. *Plant Disease*, 95 (9): 1198.
- Martin R R, Tzanetakis I E. 2006. Characterization and recent advances in detection of strawberry viruses. *Plant Disease*, 90: 384 - 396.
- Shi Y, Han X Y, Sun H, Wei Y, Wang Z Y, Li H L, Chen L L, Sun B J. 2018. First report of *Strawberry pallidosis associated virus* infecting strawberry in China. *Plant Disease*, 102 (1): 257 - 257.
- Tzanetakis I E, Reed J, Martin R R. 2005. Nucleotide sequence, genome organization and phylogenetic analysis of Strawberry pallidosis associated virus, a new member of the genus *Crinivirus*. *Archives of Virology*, 150 (2): 273 - 286.
- Wang Yun-sheng, Chen Xiao-yu, Xing Dong-mei, Ran Ce, Chen liu, Zhu Ya-dong, Shang Qiao-xia. 2015. Investigation of strawberry viruses and RT-PCR detection in some areas of China//Proceedings of the annual meeting of Chinese society for plant pathology (2015). Beijing: China Agriculture Press: 260 - 261. (in Chinese)
- 王运生, 陈笑瑜, 邢冬梅, 冉 策, 陈 柳, 朱亚东, 尚巧霞. 2015. 我国部分地区草莓病毒种类调查及 RT-PCR 检测//中国植物病理学会 2015 年学术年会. 北京: 中国农业出版社: 260 - 261.
- Yang Bo, Zhao Bao-long, Hao Xiao-jun, Du Ye-juan, He Wang. 2018. Detection and investigation of four strawberry viruses in Xinjiang region. *Xinjiang Agricultural Sciences*, 55 (9): 1689 - 1697. (in Chinese)
- 杨 波, 赵宝龙, 郝小军, 都业娟, 何 旺. 2018. 新疆地区四种草莓病毒病原的检测. *新疆农业科学*, 55 (9): 1689 - 1697.