

# 广丰千金薯烟草花叶病毒 lncRNA 测序鉴定、原核蛋白表达及其序列分析

尹明华<sup>1,2,3,4</sup>, 张艳红<sup>1</sup>, 李淑娟<sup>1</sup>, 段俊<sup>1</sup>, 蒋晓峰<sup>1</sup>

1. 上饶师范学院生命科学院, 江西 上饶 334001

2. 上饶农业技术创新研究院, 江西 上饶 334001

3. 上饶市药食同源植物资源保护与利用重点实验室, 江西 上饶 334001

4. 上饶市薯芋类作物种质保存与利用重点实验室, 江西 上饶 334001

**摘要:** 目的 对广丰千金薯 *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin 烟草花叶病毒进行鉴定、原核蛋白表达和序列分析, 为广丰千金薯脱毒苗的培育和鉴定提供理论依据。方法 通过 lncRNA 测序鉴定广丰千金薯烟草花叶病毒蛋白基因, 利用大肠杆菌重组技术表达广丰千金薯烟草花叶病毒蛋白, 并采用生物信息学方法进行序列分析。结果 广丰千金薯烟草花叶病毒蛋白基因包括复制酶 1 (ORF1)、复制酶 2 (ORF2)、RNA 聚合酶、运动蛋白 (movement protein, MP)、带电蛋白和外壳蛋白 (coat protein, CP); 广丰千金薯烟草花叶病毒 CP、MP、RNA 聚合酶、带电蛋白、ORF1 和 ORF2 基因 cDNA 总长度分别为 480、807、1425、123、4851、3351 bp, 其蛋白分别由 159、268、474、40、1613 和 1114 个氨基酸组成; 广丰千金薯烟草花叶病毒蛋白除带电蛋白为疏水性蛋白质外, 其余如 ORF1、ORF2、RNA 聚合酶、运动蛋白和外壳蛋白均为亲水性蛋白质; 广丰千金薯烟草花叶病毒带电蛋白二级结构由  $\beta$ -片层和无规则卷曲构成, CP、MP、RNA 聚合酶、ORF1 和 ORF2 的二级结构均由  $\alpha$  螺旋、 $\beta$ -片层、无规则卷曲构成; 广丰千金薯烟草花叶病毒 CP、MP、RNA 聚合酶、带电蛋白、ORF1 和 ORF2 三级结构均为单体; 广丰千金薯烟草花叶病毒 CP、MP、RNA 聚合酶、带电蛋白、ORF1 和 ORF2 的亚细胞定位分别为内质网、细胞核、细胞质、细胞核、线粒体、叶绿体; 广丰千金薯烟草花叶病毒 CP、MP、RNA 聚合酶、带电蛋白、ORF1 和 ORF2 与烟草花叶病毒的亲缘关系较近, 同源性分别为 99.79%、99.38%、99.23%、100%、99.38%、99.49%; 广丰千金薯烟草花叶病毒复制酶、RNA 聚合酶、MP 和 CP 可实现大肠杆菌重组表达, 表明广丰千金薯烟草花叶病毒的 lncRNA 测序结果准确。结论 广丰千金薯烟草花叶病毒与烟草花叶病毒同源性极高, 氨基酸序列及核酸序列与烟草花叶病毒其他植物分离物相似度高, 在进化上高度保守。

**关键词:** 广丰千金薯; 烟草花叶病毒; lncRNA 测序鉴定; 原核蛋白表达; 序列分析

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2023)11-3666-10

DOI: 10.7501/j.issn.0253-2670.2023.11.028

## lncRNA sequencing identification, protein expression and sequence analysis of tobacco mosaic virus of *Dioscorea polystachya* Guangfeng Qianjin

YIN Ming-hua<sup>1,2,3,4</sup>, ZHANG Yan-hong<sup>1</sup>, LI Shu-juan<sup>1</sup>, DUAN Jun<sup>1</sup>, JIANG Xiao-feng<sup>1</sup>

1. College of Life Sciences, Shangrao Normal University, Shangrao 334001, China

2. Institute of Shangrao Agricultural Technology Innovation Research, Shangrao 334001, China

3. Shangrao Key Laboratory of Protection and Utilization of Medicinal and Edible Plant Resources Shangrao 334001, China

4. Shangrao Key Laboratory of Germplasm Conservation and Utilization of Potato and Taro Crops, Shangrao 334001, China

**Abstract: Objective** To provide theoretical basis for the cultivation and identification of virus-free seedlings of *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin, its tobacco mosaic virus was identified, detected its protein expression and analyzed its sequence. **Methods** The protein gene of tobacco mosaic virus of *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin was identified by lncRNA sequencing, and its protein of tobacco mosaic virus was expressed by *E. coli* recombinant technology, and its gene sequence was analyzed by bioinformatics method. **Results** Tobacco mosaic virus proteins in *Dioscorea*

收稿日期: 2022-12-06

**基金项目:** 国家自然科学基金资助项目 (32060092); 国家自然科学基金资助项目 (31960079); 2022 年上饶市科技专项项目 (饶科发[2023]5 号社发类) (2022A008); 江西省科技厅重点研发计划一般项目 (20202BBG73010, 20192BBGL70050); 江西省现代农业产业技术体系建设专项 (JXARS-13-赣东站); 江西省教育厅科学技术研究项目 (GJJ211729); 上饶市科技局平台载体建设项目 (2019J017, 2020J001)

**作者简介:** 尹明华 (1973—), 女, 江西永新人, 硕士, 教授, 主要从事植物生物技术研究工作。E-mail: yinminghua04@163.com

*polystachya* Turczaninow. cv. Guangfeng Qianjin included includes replicase (ORF1), replicase (ORF2), RNA polymerase, motion protein, charged protein and coat protein; The total length of cDNA of CP, MP, RNA polymerase, charged protein, replicase (ORF1) and replicase (ORF2) genes of tobacco mosaic virus in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin was 480, 807, 1425, 123, 4851 and 3351 bp respectively, and their protein were composed of 159, 268, 474, 40, 1613 and 1114 amino acids respectively; Proteins of tobacco mosaic virus in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin, except that the charged protein was hydrophobic protein, others such as replicase (ORF1), replicase (ORF2), RNA polymerase, motion protein and coat protein were hydrophilic proteins; The secondary structure of the charged protein of tobacco mosaic virus in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin was composed of extended strand and random coil, and its secondary structures of CP, MP, RNA polymerase, replicase (ORF1) and replicase (ORF2) were composed of alpha helix, extended strand and random coil; The tertiary structures of CP, MP, RNA polymerase, charged protein, replicase (ORF1) and replicase (ORF2) of tobacco mosaic virus protein in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin were monomers; The subcellular localization of CP, MP, RNA polymerase, charged protein, replicase (ORF1) and replicase (ORF2) of tobacco mosaic virus in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin were endoplasmic reticulum, nucleus, cytoplasm, nucleus, mitochondria and chloroplast respectively; CP, MP, RNA polymerase, charged protein, replicase (ORF1) and replicase (ORF2) of tobacco mosaic virus in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin are closely related to tobacco mosaic virus, and the homologies are 99.79%, 99.38%, 99.23%, 100%, 99.38% and 99.49% respectively.; The replicase, RNA polymerase, motion protein, and coat protein of tobacco mosaic virus protein genes in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin could achieve recombinant expression of *E. coli*, indicating that its lncrna sequencing results were accurate. **Conclusion** Tobacco mosaic virus protein genes in *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin had high homology with tobacco mosaic virus, and its amino acid sequence and nucleic acid sequence were highly similar to other plant isolates of tobacco mosaic virus, which was highly conserved in evolution.

**Key words:** *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin; tobacco mosaic virus; lncRNA sequencing identification; prokaryotic protein expression; sequence analysis

广丰千金薯 *Dioscorea polystachya* Turczaninow. cv. Guangfeng Qianjin 为薯蓣科薯蓣属多年生草本植物, 原产于江西省上饶市广丰区少阳乡, 为上饶市地方特色山药品种<sup>[1]</sup>。广丰千金薯肉质根状茎可菜药兼用, 菜用绵密粉嫩、软糯爽口, 药用可健脾养胃、滋肾益精、预防心血管病、增强免疫力、延缓衰老, 是老少皆宜的滋补佳品, 经济价值高, 产品远销到东南亚和日本等国, 深受消费者青睐<sup>[2]</sup>。在生产中, 广丰千金薯由于长期采用营养繁殖, 病毒感染严重, 造成种性退化、产量下降、品质变劣、薯块变小、畸形、卖相差, 农民收入受损<sup>[3]</sup>。因此, 开展健康种苗研究, 采用生物技术的方法脱除病毒, 提纯复壮, 提高产量, 改善品质已成为广丰千金薯生产中亟待解决的问题。而脱除病毒的前提是需要鉴定广丰千金薯的病毒种类及其病毒基因序列信息。广丰千金薯的病毒种类及其基因序列信息一旦探明, 即可利用分子生物学的手段去广丰千金薯试管苗脱毒效果进行快速鉴定。目前, 对广丰千金薯的研究仅限于高产栽培<sup>[3]</sup>、微型块茎萌发<sup>[4]</sup>、离体快繁<sup>[2]</sup>、干旱胁迫<sup>[1]</sup>、遗传稳定性<sup>[5]</sup>、光合生理<sup>[6]</sup>等方面, 但广丰千金薯病毒种类及其蛋白原核表达和基因序

列分析方面的研究尚无报道。本实验通过 lncRNA 测序鉴定广丰千金薯病毒种类及其相关基因, 利用大肠杆菌重组技术表达广丰千金薯病毒蛋白, 并采用生物信息学方法对其进行序列分析, 旨在为广丰千金薯脱毒苗的培育和鉴定提供理论依据和技术基础。

## 1 材料与仪器

### 1.1 材料

样品采自江西省上饶市广信区少阳乡, 由上饶师范学院王艾平教授鉴定为千金薯 *D. polystachya* Turczaninow. cv. Guangfeng Qianjin 试管苗。

### 1.2 仪器

Illumina Hiseq 4000 测序系统 (美国 Illumina 公司)、DK-S22 型电热恒温水浴锅 (上海精宏实验设备有限公司)、GL-21M 型高速冷冻离心机 (湘仪离心机仪器有限公司)、LX-B150L 型不锈钢立式灭菌器 (合肥华泰医疗设备有限公司)、SW-CJ-1FD 型超净工作台 (苏净安泰有限公司)、ZQLY-180 型恒温摇床 (上海知楚有限公司)、JY 92-IIN 型超声破碎仪 (宁波新芝生物科技股份有限公司)、DW-86L626 型超低温冰箱 (青岛海尔特种电器有限公司)、蛋白纯化仪 (美国 Bio-rad

公司, NGC Quest™ 10 Plus)、DYCZ-24DH 型 SDS-PAGE 电泳仪(北京六一生物科技有限公司)、Qubit 3.0 型核酸蛋白定量仪(美国 Life 公司-美国生命技术公司)。

## 2 方法

### 2.1 广丰千金薯烟草花叶病毒种类及其蛋白基因的确定

**2.1.1 RNA 提取** 从江西铅山红芽芋试管苗植株上各采集 100 mg 左右幼嫩叶组织, 混合后提取小 RNA。

**2.1.2 RNA 质检** 完成 RNA 抽提后, 利用 1% 琼脂糖凝胶电泳检测 RNA 的质量。

**2.1.3 RNA 片段化** 采取约 400~500 bp 样品放入 Covaris M220 仪器中进行反转录。

**2.1.4 文库构建** cDNA 片段末端补平, 加 A, 加测序接头; 琼脂糖凝胶电泳进行片段筛选, 将目标区域的片段回收; 通过 PCR 扩增筛选未加上接头的片段; PCR 产物纯化, 质检。试剂为 TruSeq™ DNA Sample Prep Kit。

**2.1.5 桥式 PCR** DNA 片段的一端与芯片上引物碱基互补, 固定在芯片上; 另一端随机与附近的另一种引物互补, 也被固定住, 形成“桥(bridge)”; PCR 扩增, 产生 DNA 簇; 线性化成为单链。试剂为 HiSeq PE Cluster Kit v4 cBot。

**2.1.6 Illumina HiSeq 4000 测序** 加入改造过的 DNA 聚合酶和带有 4 种荧光标记的 dNTP, 每次循环只掺入单种碱基; 用激光扫描反应板表面, 读取每条模板序列第 1 轮反应所聚合上去的核苷酸种类; 将“荧光基团”和“终止基团”化学切割, 恢复 3'端黏性, 继续聚合第 2 个核苷酸; 统计每轮收集到的荧光信号结果, 获知模板 DNA 片段的序列。

利用 SOAPdenovo v2.04 (<http://soap.genomics.org.cn/>) 拼接软件对质控后的全部 clean data 进行 denovo 初步组装, 将组装得到的全部 contig 与病毒数据库进行比对, 筛选得到来自病毒的 contig; 然后利用 MITObim v1.6 通过迭代比对, 将测序的所有 clean reads mapping 到这些 contig 上进行延伸, 通过 gap close 获得病毒的全基因组序列。

利用 DOGMA 软件 (<http://dogma.cccb.utexas.edu/>) 对基因组中包含的 gene 进行预测。将预测基因的蛋白序列与 Nr 数据库进行 blastp 比对(BLAST 2.2.28+), 初步判断广丰千金薯的病毒种类及其相

关基因。

### 2.2 广丰千金薯烟草花叶病毒蛋白基因的生物信息学分析

使用 BioEdit 软件翻译基因序列为氨基酸序列, 用 ProtParam 预测酶的理化性质, 用 ProtScale 预测酶的疏/亲水性。使用 GOR I 软件在线预测酶的二级结构。使用 SWISS-MOLD 在线预测酶的三级结构。采用 WoLFPSort 在线预测基因的表达部位。通过软件 DNAMAN 和 Bioedit 进行氨基酸序列比对, 利用 MEGA5.0 软件进行系统进化树的构建。

### 2.3 广丰千金薯烟草花叶病毒蛋白的大肠杆菌重组表达

**2.3.1 感受态转化及阳性克隆筛选** 从超低温冰箱中取出 BL21 (DE3) 感受态细胞, 置于冰上融化; 加入质粒 (pET-28a, 5 μg), 轻轻吹吸充分混匀, 冰上放置 30 min; 水浴锅 42 °C 热击 90 s, 冰上放置 1 min; 加入 800 μL 预热 LB 液体培养基, 37 °C 158 r/min 培养 50 min; 6000 r/min 离心 4 min, 去部分上清 (800 μL 体积), 剩余菌液混匀后涂至氨苄抗性平板上; 倒置平板, 37 °C 培养 12~16 h 可见出现菌落 (阳性克隆)。

**2.3.2 阳性克隆小量表达与鉴定** 挑选含重组质粒的单菌落至 3 mL LB 液体培养基 (氨苄抗性) 中, 37 °C 培养过夜后 -20 °C 保种; 分别挑选含重组质粒的单菌落至 3 mL LB 液体培养基 (氨苄抗性) 中, 37 °C 震荡培养至 A<sub>600</sub> 约 0.6; 取部分菌液作为对照组, 余下菌液加入 IPTG 诱导剂 (终浓度 0.1 mmol/L), 16 °C 震荡培养 12 h; 分别取 2 组菌液 0.15 mL, 12 000×g 离心 2 min, 菌体沉淀以 40 μL 1×loading buffer 重悬裂解, SDS-PAGE 检测。

**2.3.3 蛋白大量表达及破菌检测** 取保存于 -20 °C 的菌种 100 μL 接种于 100 mL LB 液体培养基 (氨苄抗性) 中震荡培养过夜; 取 100 mL 菌液接种于 2000 mL LB 液体培养基中, 37 °C 扩大培养至 A<sub>600</sub> 约 0.6, 降低培养温度到 30 °C; 加入 IPTG 诱导剂至终浓度 0.1 mmol/L, 16 °C 继续震荡培养 12 h; 8000 r/min 离心 3 min 收集菌体, 重悬于 50 mL 预冷 NTA-0 缓冲液中, 冰浴 30 min; 超声破碎菌体, 参数设置为功率 200 W、工作 3 s、暂停 4 s、99 个循环; 16 000 r/min, 4 °C 离心 50 min, 收集上清以及沉淀; 取少量上清及沉淀进行 SDS-PAGE 检测, 剩余上清及沉淀至于 4 °C 备用。

### 3 结果与分析

#### 3.1 测序数据统计

对经过质量剪切前后的数据分别进行测序总碱基数、文库平均插入长度、 $Q_{20}$ 、 $Q_{30}$ 、平均测序深度统计,结果见表 1。由表 1 可知,建立的文库准确率很高。

表 1 测序数据统计

Table 1 Sequencing data statistics

| 统计项目          | 正向测序               | 反向测序               | 2 次测序总和               |
|---------------|--------------------|--------------------|-----------------------|
| 总数            | 45 481 481         | 45 481 481         | 90 962 962            |
| 总长度/bp        | $6.82 \times 10^9$ | $6.82 \times 10^9$ | $1.36 \times 10^{10}$ |
| 平均 Read 长度/bp | 150                | 150                | 150                   |
| $Q_{10}$ /bp  | $6.82 \times 10^9$ | $6.82 \times 10^9$ | $1.36 \times 10^{10}$ |
| $Q_{10}$ 比率/% | 100.00             | 100.00             | 100.00                |
| $Q_{20}$ /bp  | $6.63 \times 10^9$ | $6.63 \times 10^9$ | $1.33 \times 10^{10}$ |
| $Q_{20}$ 比率/% | 97.25              | 97.19              | 97.22                 |
| $Q_{30}$ /bp  | $6.37 \times 10^9$ | $6.33 \times 10^9$ | $1.27 \times 10^{10}$ |
| $Q_{30}$ 比率/% | 93.42              | 92.78              | 93.10                 |
| $N$ /bp       | 90 971             | 688                | 91 659                |
| $N$ 比率/%      | 0.00               | 0.00               | 0.00                  |
| GC/bp         | $3.82 \times 10^9$ | $3.8 \times 10^9$  | $7.62 \times 10^9$    |
| GC 比率/%       | 55.96              | 55.76              | 55.86                 |

#### 3.2 基因组组装

利用 SOAPdenovo v2.04 拼接软件对质控后的全部 clean data 进行 denovo 初步组装,将组装得到的全部 contig 与病毒数据库进行比对,筛选得到来自病毒的 contig; 然后利用 MITObim v1.6 通过迭代比对,将测序的所有 clean reads mapping 到这些 contig 上进行延伸,通过 gap close 获得病毒的全基因组序列,最终组装结果的统计,广丰千金薯病毒种类是 1 个,该病毒的总长度为 6395 bp,其 GC 含量为 43.35%。

#### 3.3 Gene 查找

利用 DOGMA 软件对基因组中包含的 gene 进行预测,然后在进行人工矫正。gene 预测结果的统计可知,基因数量为 6 个,基因总长度为 6264 bp,基因平均长度为 1043,基因间区 GC 含量为 44.65%。

#### 3.4 基因功能注释

将预测基因的蛋白序列与 Nr 数据库进行 blastp 比对 (BLAST 2.2.28+),从而获得预测基因的注释信息。注释结果如表 2 所示。由表 2 可知,广丰千金薯烟草花叶病毒蛋白基因包括复制酶 1 (replicase,

表 2 基因功能注释结果

Table 2 Gene function annotation results

| 基因名称 | 链 | 起始   | 终止   | 长度/aa      | 注释结果    |
|------|---|------|------|------------|---------|
| ORF1 | + | 69   | 4919 | 48 511 616 | 复制酶     |
| ORF2 | + | 69   | 3419 | 33 511 116 | 复制酶     |
| ORF3 | + | 3495 | 4919 | 1425 474   | RNA 聚合酶 |
| ORF4 | + | 4903 | 5709 | 807 268    | 运动蛋白    |
| ORF5 | + | 5666 | 5788 | 123 40     | 带电蛋白    |
| ORF6 | + | 5712 | 6191 | 480 159    | 外壳蛋白    |

ORF1)、复制酶 2 (ORF2)、RNA 聚合酶 (RNA polymerase)、运动蛋白 (movement protein, MP)、带电蛋白 (charged protei) 和外壳蛋白 (coat protein, CP)。

#### 3.5 广丰千金薯烟草花叶病毒基因 cDNA 序列

广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 基因 cDNA 总长度和 (G+C) 含量见表 3。广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 基因序列结果已经上传至 GenBank 数据库,基因登陆号为 OL944011。

表 3 广丰千金薯烟草花叶病毒基因 cDNA 序列

Table 3 cDNA sequence of tobacco mosaic virus gene from *D. polystachya* Guangfeng Qianjin

| 基因             | cDNA 序列总长度/bp | (G+C)/% |
|----------------|---------------|---------|
| ORF6 CP 基因     | 480           | 43.54   |
| ORF4 MP 基因     | 807           | 41.39   |
| ORF3 RNA 聚合酶基因 | 1425          | 40.28   |
| ORF5 带电蛋白基因    | 123           | 43.09   |
| ORF1 复制酶基因     | 4851          | 43.45   |
| ORF2 复制酶基因     | 3351          | 44.88   |

#### 3.6 广丰千金薯烟草花叶病毒蛋白氨基酸序列

Protparam 预测显示广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 氨基酸序列分析见表 4。由表 4 可知,广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 分别由 159、268、474、40、1613 和 1114 个氨基酸组成,相对分子质量分别为 17 649.81、29 981.75、54 573.87、4 845.84、183 048.95 和 125 860.14,等电点分别为 5.09、8.90、6.83、11.89、6.57、6.53。

表 4 广丰千金薯烟草花叶病毒蛋白氨基酸序列  
Table 4 Amino acid sequence of tobacco mosaic virus protein from *D. polystachya* Guangfeng Qianjin

| 蛋白           | 氨基酸数量 | 相对分子质量 | 等电点   |
|--------------|-------|--------|-------|
| ORF6 CP      | 480   | 43.54  | 5.09  |
| ORF4 MP      | 807   | 41.39  | 8.90  |
| ORF3 RNA 聚合酶 | 1425  | 40.28  | 6.83  |
| ORF5 带电蛋白    | 123   | 43.09  | 11.89 |
| ORF1 复制酶     | 4851  | 43.45  | 6.57  |
| ORF2 复制酶     | 3351  | 44.88  | 6.53  |

3.7 广丰千金薯烟草花叶病毒蛋白亲疏水性分析  
经 ProtScale 软件分析, 广丰千金薯烟草花叶病毒蛋白除 ORF5 为疏水性蛋白质外, 其余如 ORF6、ORF4、ORF3、ORF1 和 ORF2 均为亲水性蛋白质。

3.8 广丰千金薯烟草花叶病毒蛋白二级结构分析  
通过 GOR 软件预测, 广丰千金薯烟草花叶病毒蛋白二级结构分析见表 5。由表 5 可知, 广丰千金薯烟草花叶病毒 ORF5 二级结构由  $\beta$ -片层和无规

表 5 广丰千金薯烟草花叶病毒蛋白二级结构  
Table 5 Secondary structures of tobacco mosaic virus proteins from *D. polystachya* Guangfeng Qianjin

| 蛋白           | $\alpha$ 螺旋/% | $\beta$ -片层/% | 无规则卷曲/% |
|--------------|---------------|---------------|---------|
| ORF6 CP      | 37.11         | 15.09         | 47.80   |
| ORF4 MP      | 32.09         | 23.88         | 44.03   |
| ORF3 RNA 聚合酶 | 34.81         | 18.78         | 46.41   |
| ORF5 带电蛋白    | 0.00          | 47.50         | 52.50   |
| ORF1 复制酶     | 39.12         | 18.54         | 42.34   |
| ORF2 复制酶     | 41.47         | 18.13         | 40.39   |

则卷曲构成, ORF6、ORF4、ORF3、ORF1 和 ORF2 的二级结构均由  $\alpha$  螺旋、 $\beta$ -片层、无规则卷曲构成。

3.9 广丰千金薯烟草花叶病毒蛋白三级结构分析  
SWISS-MODEL 预测显示广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 三级结构均为单体 (图 1)。

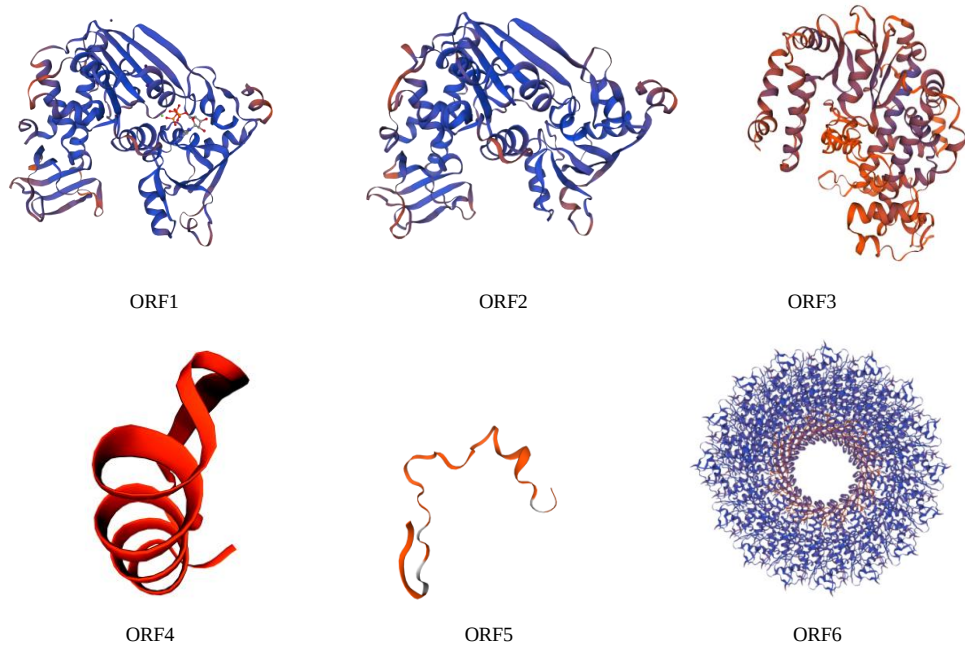


图 1 广丰千金薯烟草花叶病毒蛋白三级结构

Fig. 1 Tertiary structure of proteins of tobacco mosaic virus in *D. polystachya* Guangfeng Qianjin

3.10 广丰千金薯烟草花叶病毒蛋白亚细胞定位  
采用 WoLFPSort 在线软件对广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 基因的表达部位进行预测, 结果表明, 广丰千金薯烟草花叶病毒 ORF1、ORF2、ORF3、ORF4、ORF5 和 ORF6 分别定位于内质网、细胞核、细胞质、细

胞核、线粒体、叶绿体中。  
3.11 广丰千金薯烟草花叶病毒蛋白系统进化分析  
从构建的进化树 (图 2) 中可见, 广丰千金薯烟草花叶病毒 ORF6 与烟草花叶病毒 CP 在一个分支下, 这说明广丰千金薯烟草花叶病毒 ORF6 在进化上与烟草花叶病毒 CP 的亲缘关系较近, 同源性

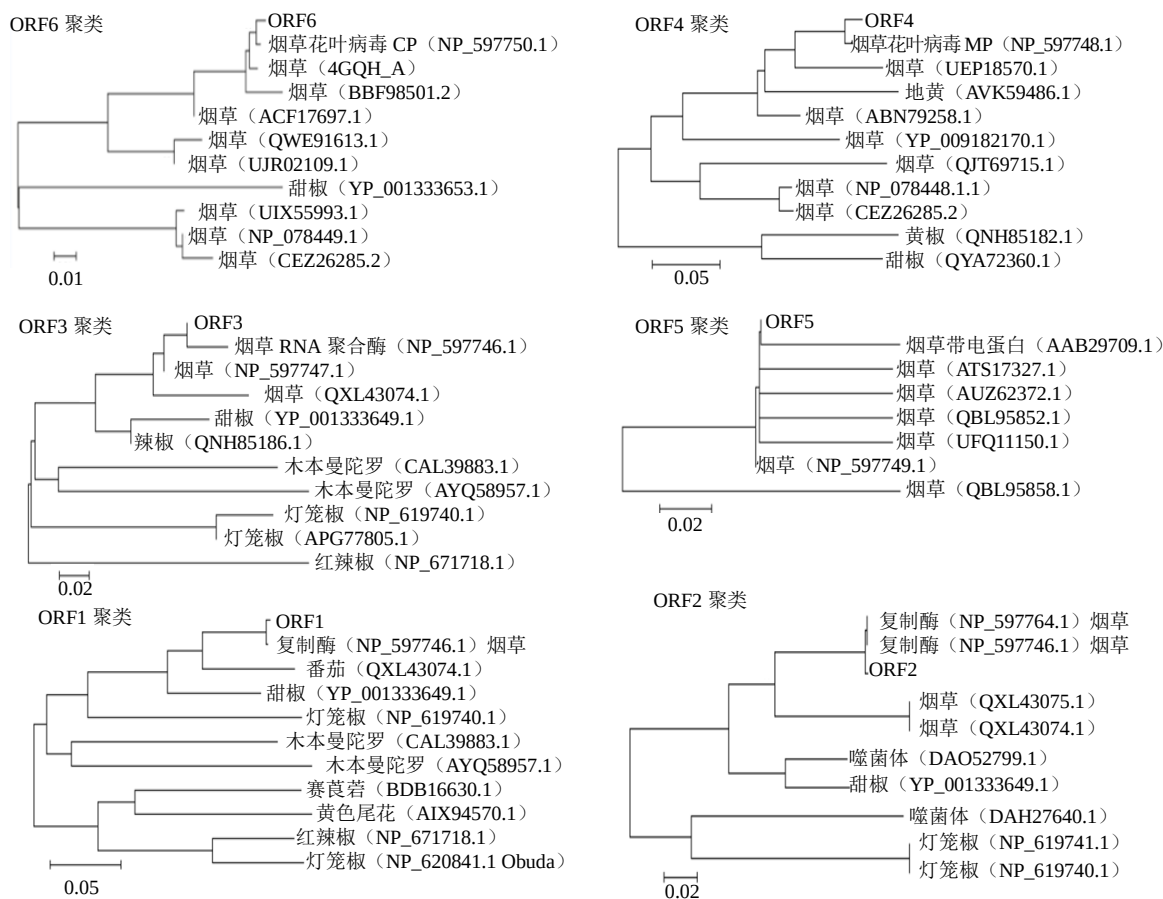


图 2 广丰千金薯烟草花叶病毒蛋白系统进化分析

Fig. 2 Phylogenetic analysis of proteins of tobacco mosaic virus in *D. polystachya* Guangfeng Qianjin

为 99.79%；广丰千金薯烟草花叶病毒 ORF4 与烟草花叶病毒 MP 在一个分支下，这说明广丰千金薯烟草花叶病毒 ORF4 在进化上与烟草花叶病毒 MP 的亲缘关系较近，同源性的 99.38%；广丰千金薯烟草花叶病毒 ORF3 与烟草花叶病毒 RNA 聚合酶在一个分支下，这说明广丰千金薯烟草花叶病毒 RNA 聚合酶 (ORF3) 在进化上与烟草花叶病毒 RNA 聚合酶的亲缘关系较近，同源性的 99.23%；广丰千金薯烟草花叶病毒带电蛋白 (ORF5) 与烟草花叶病毒带电蛋白在一个分支下，这说明广丰千金薯烟草花叶病毒带电蛋白 (ORF5) 在进化上与烟草花叶病毒带电蛋白的亲缘关系较近，同源性的 100%。广丰千金薯烟草花叶病毒复制酶 (ORF1) 与烟草花叶病毒复制酶在一个分支下，这说明广丰千金薯烟草花叶病毒复制酶 (ORF1) 在进化上与烟草花叶病毒复制酶的亲缘关系较近，同源性的 99.38%。广丰千金薯烟草花叶病毒复制酶 (ORF2) 与烟草花叶病毒复制酶在一个分支下，这说明广丰千金薯烟草花叶病

毒复制酶 (ORF2) 在进化上与烟草花叶病毒复制酶的亲缘关系较近，同源性的 99.49%。

### 3.12 广丰千金薯烟草花叶病毒蛋白的序列比对信息

广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 同源蛋白的序列比对信息见图 3。图 3 中“\*”号区域是该蛋白家族的保守结构域。从图 3 可知，广丰千金薯烟草花叶病毒 ORF6、ORF4、ORF3、ORF5、ORF1 和 ORF2 的序列与烟草花叶病毒相似，同源性均在 99.49% 以上，因此，可以利用烟草花叶病毒基因序列来进行广丰千金薯烟草花叶病毒的种类鉴别。

### 3.13 广丰千金薯烟草花叶病毒蛋白的大肠杆菌重组表达

转化 BL21 (DE3) 感受态细胞，挑选 AMP 抗性平板上长出的阳性斑进行扩大培养与表达纯化，结果见图 4。由图 4 可知，除了 ORF5 蛋白太小 (485 000) 无法正常表达外，其他 5 个蛋白 (ORF1、ORF2、ORF3、ORF4 和 ORF6) 都成功表达，蛋白大小分别为 183 050、125 860、54 570、29 980 和 17 650。



NP\_597749.1 MMIRRRLLSPNRIKPKYVLYQHYHSISVRVLVHSVGRPNRNV  
 AAB29709.1 -MIRRRLLPIRIIRFKYVLYQHYHSISVRVLVHSVGRPNRNV  
 ORF5 MMIRRRLLSPNRIKPKYVLYQHYHSISVRVLVHSVGRPNRNV  
 QBL95852.1 MMIRRRLLSPNRIKPKYVLYQHYHSISVRVLVHSVGRPNRNV  
 AT517327.1 MMIRRRLLSPNRIKPKYVLYQHYHSISVRVLVHSVGRPNRNV  
 UFQ11150.1 MMNRRLSPNRIKPKYVLYQHYHSISVRVLVHSVGRPNRNV  
 AUZ62372.1 MMIRRRLLSPNRIKPKYVLYQHYHSISVRVLVHSVGRPNRNV  
 QBL95858.1 MMIRRRLLSPNRIKPKYVLYQHYHSISVRVLVHSVGRPDNV  
 1 10 20 30 40

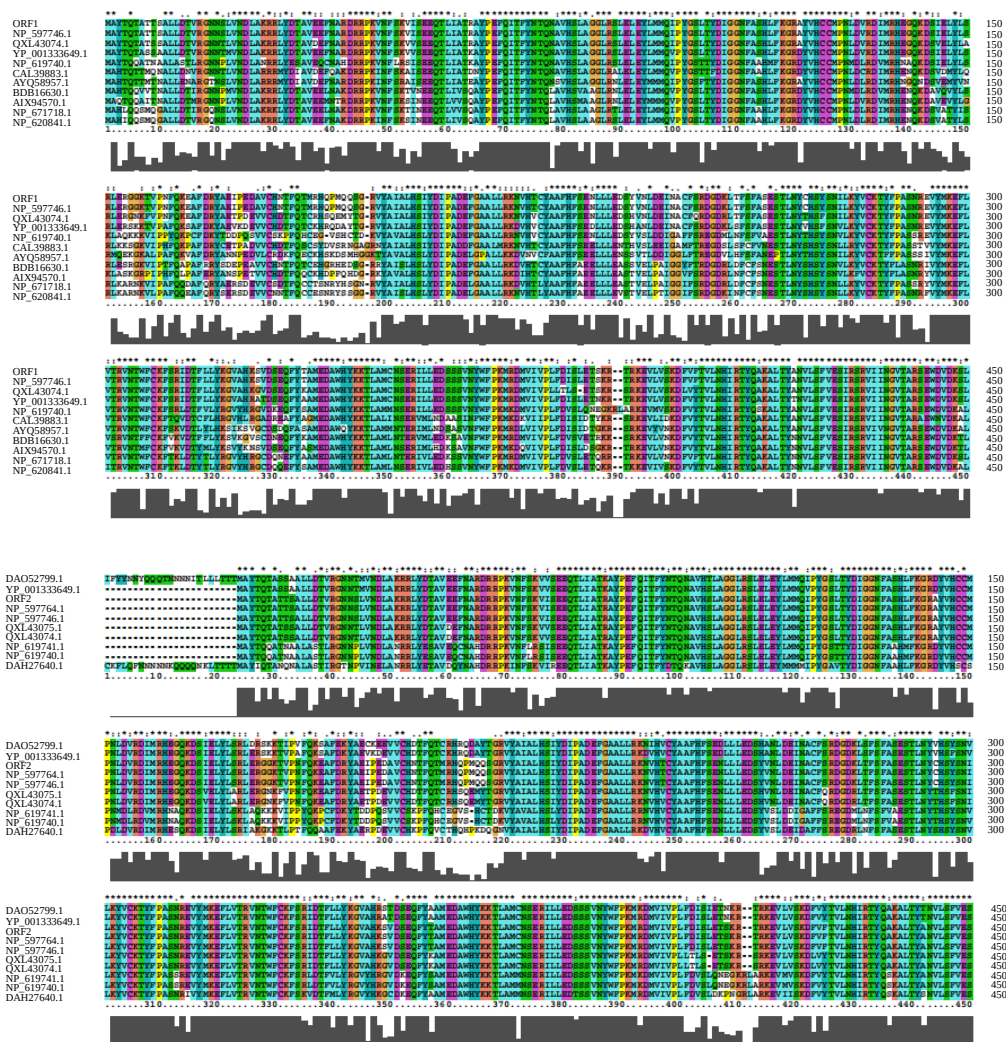


图3 广丰千金薯烟草花叶病毒蛋白氨基酸序列的同源性比较

Fig. 3 Homology comparison of amino acid sequences of proteins of tobacco mosaic virus in *D. polystachya* Guangfeng Qianjin

这说明, 广丰千金薯烟草花叶病毒蛋白基因的 IncRNA 测序结果是准确的。

#### 4 讨论

目前报道的山药病毒病害主要为山药花叶病毒病, 生长初期为害症状不明显, 到生长中后期症状严重, 叶上出现褪绿斑, 随后绿色花斑隆起, 呈斑驳花叶、凸凹卷曲畸形、黄化或坏死, 严重时整株矮化, 生长缓慢, 地下茎缩小, 畸形, 产量下降<sup>[7]</sup>。鄢明峰等<sup>[8]</sup>发现日本山药花叶病毒 (Japanese yam mosaic virus, JYMV) 为瑞昌山药花叶型病毒的病毒原, 江西瑞昌、广西和日本共 32 个 JYMV 分离物按其地域来源形成了 3 个独立的进化组群, 表明 JYMV 的进化与其地理来源具有明显的相关性。张蕾等<sup>[9]</sup>和 Zou 等<sup>[10]</sup>均利用 RT-PCR 手段检测到淮山

药感染山药温和花叶病毒 (Yam mild mosaic virus, YMMV)。Silva 等<sup>[11]</sup>利用 RT-PCR 技术确定西非山药的主要危害病毒为山药花叶病毒。在本试验中, 通过 IncRNA 测序鉴定广丰千金薯烟草花叶病毒基因, 并利用大肠杆菌重组技术成功表达了广丰千金薯烟草花叶病毒蛋白, 第 1 次确定了广丰千金薯的主要危害病毒为烟草花叶病毒。

研究表明, 山药花叶病毒 (yam mosaic virus, YMV) 是一种具有大约 785 nm 长的弯曲丝状颗粒, 由一种单链 RNA 和一种外壳蛋白组成, 很难传播到小范围的宿主, 由蚜虫以非持久性方式传播, 在热带地区的山药种植中造成了重要的经济损失<sup>[12]</sup>。YMV 基因组 RNA 长度为 9608 个核苷酸, 包含编码 3103 个氨基酸 (aa) 多蛋白的一个开放阅读框



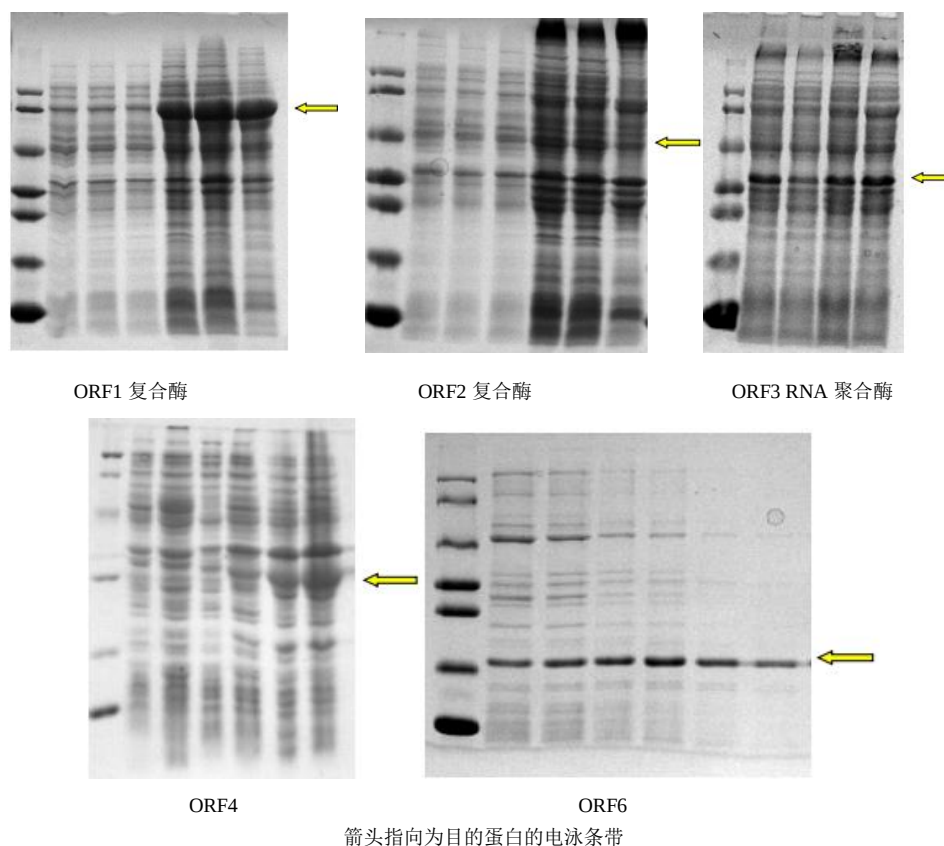


图 4 广丰千金薯烟草花叶病毒的蛋白表达

Fig. 4 Protein expression of tobacco mosaic virus in *D. polystachya* Guangfeng Qianjin

(open reading frame, ORF) [13], 具有马铃薯 Y 病毒属特征, 是马铃薯 Y 病毒属的一个独特成员 [14]。史跃伟等 [15] 通过 RT-PCR 的方法获得了贵州烟草花叶病毒 CP 基因, 该基因全长为 507 bp, 并利用原核表达技术表达了贵州烟草花叶病毒 CP 蛋白, 为后期单克隆抗体的制备奠定了基础。烟草花叶病毒南瓜分离物 CP 基因与 GenBank 上其他 TMV 分离物 CP 基因核苷酸序列的同源性为 86.5%~99.0%, 推导的氨基酸序列同源性为 93.7%~99.4%, 烟草花叶病毒南瓜分离物 CP 基因与原核表达载体 pET-22b (+) 连接, 在大肠杆菌 BL21 (DE3) pLysS 诱导表达出相对分子质量约 20 000 的融合蛋白 [15]。云南 5 个不同地区烟草花叶病毒 CP 基因均为 480 个碱基, 编码 159 个氨基酸, 与 TMV-U1、P、B 株和福建分离物有较高的同源率 (90% 以上), 而与 TMV 中国株系的同源率较低 (70%~80%), 与 TMV 中国株应属于同种病毒不同的株系 [16]。李凡等 [17] 的研究结果也证实了这一观点, 再次确认烟草花叶病毒云南分离物 CP 基因为 480 个碱基, 编码 159 个氨基酸, 并指出烟草花叶病毒云南分离物 CP 基因与 T

MV-U1 株系和 TMV 韩国普通株系核苷酸同源性均为 100%。但也有研究表明, 云南烟草花叶病毒 CP 基因含 477 个核苷酸, 编码 159 个氨基酸 [18]。本实验结果也表明, 广丰千金薯烟草花叶病毒 CP 基因 cDNA 总长度为 480 bp, 编码 159 个氨基酸组成, 与烟草花叶病毒 CP 基因亲缘关系较近。

烟草花叶病毒的致病性是由 CP 基因及 3' 端非编码区以外的基因 (如 MP 基因、复制酶基因等) 决定的。烟草花叶病毒 MP 基因决定植物病毒在寄主植物细胞间的移动从而可决定病毒对寄主的侵染性。有研究表明, 烟草花叶病毒蚕豆分离 MP 基因由 807 个核苷酸组成, 编码 268 个氨基酸, 烟草花叶病毒的 MP 基因在不同株系中非常保守, 同源性极高 [19]。王得元等 [20] 对广东烟草花叶病毒 MP 基因进行 PCR 扩增, 获得了一条 0.8 kb 的特异扩增产物, 与已发表的 TMV MP 基因编码区 803 bp 大小相近。余晓红等 [21] 克隆了烟草花叶病毒 MP 基因, 发现 MP 基因含有 1 个 507 bp 的阅读框架, 编码 268 个氨基酸, 与国外发表的 TMVU1 株运动蛋白基因 cDNA 的核苷酸序列及推导的氨基酸序列相比, 同

源率均高达 98 以上,再次表明烟草花叶病毒 MP 基因在同一病毒的不同株系中是相当保守的。本实验结果表明,广丰千金薯烟草花叶病毒 MP 基因的 cDNA 总长度为 807 bp,编码 268 个氨基酸,与烟草花叶病毒 MP 基因亲缘关系较近。

因此,广丰千金薯烟草花叶病毒与烟草花叶病毒的亲缘关系较近,同源性在 99%以上。表明广丰千金薯烟草花叶病毒不属于马铃薯 Y 病毒属,与烟草花叶病毒同源性极高,氨基酸序列及核酸序列与烟草花叶病毒其他植物分离物相似度高,在进化上高度保守。因此,可以利用烟草花叶病毒基因序列来进行广丰千金薯烟草花叶病毒的种类鉴别。

**利益冲突** 所有作者均声明不存在利益冲突

#### 参考文献

- [1] 柯维忠,王丽,杨星鹏,等.广丰千金薯驯化移栽苗对 PEG 干旱胁迫的光合生理响应 [J]. 浙江农业学报, 2016, 28(9): 1462-1475.
- [2] 尹明华,徐志坚,章省琴,等.广丰千金薯离体快繁及其气孔观察、染色体倍数 FCM 分析和 DNA 变异 ISSR 检测 [J]. 中药材, 2016, 39(7): 1446-1451.
- [3] 邱翠金,何鸿弘,叶为波,等.广丰千金薯高产栽培主要技术措施 [J]. 现代园艺, 2009(6): 27, 31.
- [4] 尹明华,王丽,徐玉琴,等.江西产山药微型块茎萌发苗耐盐性隶属函数及主成分分析 [J]. 中草药, 2016, 47(14): 2526-2533.
- [5] 尹明华,徐志坚,黄玮,等.江西山药种质资源遗传多样性及其组培苗遗传稳定性的 RAPD 检测 [J]. 中草药, 2016, 47(19): 3486-3493.
- [6] 洪森荣,王丽,杨星鹏,等.江西地方特色山药盆栽试管苗叶绿素含量以及其光合荧光参数的变化 [J]. 基因组学与应用生物学, 2017, 36(10): 4357-4366.
- [7] 黄婷,蒋军喜,余国庆,等.山药病害最新研究进展 [J]. 生物灾害科学, 2014, 37(1): 74-78.
- [8] 鄢明峰,贺哲,黄婷,等.瑞昌山药花叶型病毒病毒原的分子鉴定及其遗传变异性分析 [J]. 江西农业大学学报, 2017, 39(6): 1126-1130.
- [9] 张蕾,邹承武,蒙姣荣,等.抗病毒药剂处理脱除淮山药温和花叶病毒试验 [J]. 南方农业学报, 2014, 45(12): 2163-2166.
- [10] Zou C, Meng J, Li Z, et al. First report of yam mild mosaic virus in yam in Guangxi province, China [J]. *Plant Dis*, 2011, 95(10): 1320.
- [11] Silva G, Bömer M, Nkere C, et al. Rapid and specific detection of yam mosaic virus by reverse-transcription recombinase polymerase amplification [J]. *J Virol Methods*, 2015, 222: 138-144.
- [12] Duterme O, Colinet D, Kummert J, et al. Determination of the taxonomic position and characterization of yam mosaic virus isolates based on sequence data of the 5'-terminal part of the coat protein cistron [J]. *Arch Virol*, 1996, 141(6): 1067-1075.
- [13] Aleman M E, Marcos J F, Brugidou C, et al. The complete nucleotide sequence of yam mosaic virus (Ivory Coast isolate) genomic RNA [J]. *Arch Virol*, 1996, 141(7): 1259-1278.
- [14] Duterme O, Colinet D, Kummert J, et al. Determination of the taxonomic position and characterization of yam mosaic virus isolates based on sequence data of the 5'-terminal part of the coat protein cistron [J]. *Arch Virol*, 1996, 141(6): 1067-1075.
- [15] 史跃伟,郭玉双.烟草花叶病毒贵州分离物 CP 基因的原核表达及纯化 [J]. 贵州农业科学, 2012, 40(12): 18-20.
- [16] 丁铭,方琦,张丽珍,等.云南五个不同地区烟草花叶病毒外壳蛋白基因的序列比较 [J]. 中国病毒学, 2004, 19(1): 76-78.
- [17] 李凡,周雪平,陈海如,等.云南五个不同地区烟草花叶病毒外壳蛋白基因的序列比较 [J]. 浙江大学学报: 农业与生命科学版, 2000, 26(3): 261-265.
- [18] 王芳,鄢波,王玲仙,等.云南烟草花叶病毒外壳蛋白基因的分离克隆及序列分析 [J]. 西南农业学报, 1999, 12(2): 7-11.
- [19] 陈青,周雪平,薛朝阳,等.烟草花叶病毒蚕豆分离物移动蛋白基因克隆及序列分析 [J]. 浙江农业大学学报, 1999, 25(3): 239-242.
- [20] 王得元,刘志昕,潘俊松,等.广东烟草花叶病毒移动蛋白基因的 cDNA 克隆和酶切鉴定 [J]. 江西农业大学学报, 1998, 20(3): 334-338.
- [21] 余晓红,朱玉贤,应华澄,等.烟草花叶病毒运动蛋白基因的 cDNA 克隆、序列测定及植物转化 [J]. 生物工程学报, 1996, 12(2): 211-214.
- [22] 余晓红,朱玉贤,应华澄,等.烟草花叶病毒运动蛋白基因的 cDNA 克隆、序列测定及植物转化 [J]. 生物工程学报, 1996, 12(2): 211-214.

[责任编辑 时圣明]