

马铃薯卷叶病毒中国株(PLRV-Ch)ORF2a 基因特征分析

赵国芬² 张鹤龄^{1*} 哈斯阿古拉¹

(¹内蒙古大学 呼和浩特 010021, ²内蒙古农业大学 呼和浩特 010018)

关键词 马铃薯卷叶病毒, 基因克隆, 序列分析, ORF2a

中图分类号 Q78 文献标识码 A 文章编号 1000-3061(2002)06-0744-05

马铃薯卷叶病毒(PLRV)是正链 RNA 病毒,属黄化病毒组^[1]严格虫传,分布广泛,难以控制,侵染马铃薯,给生产造成巨大损失。PLRV 基因全长 6.0 kb,有 6 个读码框架,其中 ORF2a 是第二读框,全长 1920bp,编码一个 70kD 的多肽。另外,ORF2a 在与 ORF2b 重叠处可发生移码继续转译,直到 ORF2b 的尾,转译产物为一条 118kD 的多肽,该蛋白的 C 端与复制酶的序列具很大的同源性^[2-4]。

Prufer^[5]等和 Kujawa^[6]等分别研究了 PLRV 基因组上 ORF2a 与 ORF2b 重叠区附近与移码有关的滑动序列及其下游的茎环结构。Gorbalenya 等研究了黄化病毒组 2 的 ORF2a 编码的蛋白,发现所有的成员都有类似于糜蛋白酶的特征序列^[7]。Goldbach 等研究了正链 RNA 病毒的螺旋酶的特征序列,发现 PLRV 具有 motif IV 与 motif VI。Prufer 等发现 PLRV 的 ORF2a 超读产物有一个富含精氨酸的区域^[8]。

自 1991 年 Golemboski 等利用 TMV 复制酶基因编码的 54kD 蛋白的基因转化烟草获得一定抗病性^[9],国内近几年在复制酶介导的抗性方面取得了很大进展。但是,尚未见有关 ORF2a 介导的抗性报道。

从 1992 年起,本室陆续报道了马铃薯卷叶病毒的外壳蛋白基因克隆和序列分析^[10]、56kD 蛋白基因克隆和序列分析^[11]、基因间隔区克隆和序列分析^[12]、复制酶基因 3'端的克隆和序列分析^[13]等。本文首次报道 PLRV 中国株的 ORF2a 的基因全序列,分析了其基因结构特点及所编码的氨基酸特征,并与其它株系进行了比较,为其功能的研究创造了条件。

1 材料与方 法

1.1 试剂和酶

反转录试剂盒购自美国华美生物工程公司。T4DNA 连接酶 X-gal IPTG 等购自美国 Promega 公司,限制性内切酶购自德国 Boehringer Mannheim 公司或华美生物工程公司。

1.2 菌种及质粒

E. coli JM109 为内蒙古大学微生物教研室保存, pUC19 购自华美生物工程公司。

1.3 病毒

马铃薯卷叶病毒中国株由内蒙古大学微生物教研室的分离自“紫花白”马铃薯并保存。

1.4 病毒提纯及 RNA 提取

病毒提纯按内蒙古大学微生物教研室报道的方法^[14]进行, RNA 提取用 SDS-酚法。

1.5 引物的设计

根据国外已报道的 PLRV 四个株系的全序列设计合成了下列几个引物:ORF2a 的 3'端片段(长 1.0 kb,相应于 PLRV-S 从 1151-2234 核苷酸)的两引物:上游引物 P5, 31 mer 5'-CAGGATCCAGTGCCCGTAATGATATCTCCAT-3', 下游引物 P1, 28mer, 5'-AGGGTACCTGTAGATTCAGGCTTTGGAG-3'。ORF2a 的 5'端片段(长 0.6 kb 相应于 PLRV-S 从 307-918 核苷酸)的两引物:上游引物 P2, 34 mer, 5'-CGGGATCCAACCAT-GAACAGATTTACCGCATATG-3', 下游引物 P7, 28mer 5'-CAGG-TACCTTCACAG-CCTTCTCATTCTT-3'。ORF2a 中间片段(长 0.4 kb 相应于 PLRV-S 从 874-1258 核苷酸)的两引物:上游引物 P874, 31 mer, 5'-CAGGATCCGATTATTTGGAGGGCAGTTTC-3'下游引物 P6, 28mer, 5'-ATGGTACCTCAAGAGTGTAGAAAGAGGC-3'。P5、P2、P874 的 5'端设计有 BamH I 酶切位点, P1、P6、P7 的 5'端设计有 Kpn I 酶切位点。

1.6 cDNA 第一链合成

分别利用特异性 P1、P6、P7 以 PLRV-Ch RNA 为模板,按反转录试剂盒提供的方法合成 cDNA 的第一链。

1.7 PCR 扩增

分别以各自 cDNA 第一链为模板进行 PCR,扩增体系为 50 μ L 无菌水 29 μ L, 4 \times dNTP(2.5mmol/L) 5 μ L, 10 \times Buffer 5 μ L, 引物各 5 μ L(5pmol), 模板 1 μ L, 石蜡油 40 μ L。P1、P5 为一对引物, 95 $^{\circ}$ C 5min, 循环条件为 95 $^{\circ}$ C 变性 1min, 50 $^{\circ}$ C 退火 1min,

收稿日期 2002-05-31, 修回日期 2002-08-26。

基金项目 国家自然科学基金资助项目(No. 9270453)。

* 通讯作者。 Tel 086-471-4992693; Fax 86-471-4992442; E-mail hlzhang@nmg2.imu.edu.cn

72℃延伸 2min,30 个循环后 72℃延伸 10min,为产物 II;P2、P7 为一对引物,95℃5min,循环条件为 95℃变性 1min,50℃退火 1min,72℃延伸 2min,30 个循环后 72℃延伸 10min,为产物 I;P6、P874 为一对引物,94℃4min,循环条件为 94℃变性 1min,60℃退火 1min,72℃延伸 2min,30 个循环后 72℃延伸 10min,为产物 III。PCR 扩增后电泳检测,观察结果。

1.8 cDNA 克隆

PCR 产物 I、II 采用酚:氯仿:异戊醇(25:24:1)抽提,乙醇沉淀的方法进行纯化,用低熔点胶法回收。*Bam*H I、*Kpn* I 双酶切,与 *Bam*H I、*Kpn* I 双酶切 pUC19 连接,转化至 JM109 中,在含 X-gal、Amp、IPTG 的平皿上培养,挑选白色菌落进行筛选。PCR 产物 I、II 获得的克隆分别为 pLR8、pLR7。

1.9 重组质粒鉴定

碱法提取质粒,分别用 PCR 方法、不同限制内切酶酶切分析。

1.10 序列分析

将 pLR8、pLR7 送北京赛百盛公司用自动荧光测序仪进行序列分析,pLR7 从两端各进行了一次,pLR8 进行了一次分析;PCR 产物 III 由上海生工 ABI PRISM377DNA sequencer-B 进行的序列分析。

2 结果

2.1 目的片段 RT-PCR 扩增及克隆

先以 PLRV-Ch RNA 为模板,分别以引物 P1、P6、P7 反转录合成 cDNA 的第一条链分别以三对引模板进行 PCR 扩增并对三个产物 I、II、III 进行了电泳检测,产物 I、II、III 的长度分别为 0.6 kb、1kb、0.4 kb 结果表明得到了预期大小的片段。将 PCR 产物 I、II 分别克隆于 pUC19 的 *Bam*H I、*Kpn* I 酶切位点,得到重组质粒 pLR8、pLR7,转化大肠杆菌 JM109 中,在含 X-gal、Amp、IPTG 的平皿上培养,挑选白色菌落进行筛选与鉴定,证明得到了预期的两个克隆。

2.2 核苷酸序列分析

核苷酸序列分析结果(见图 1)表明中国株 PLRV-Ch ORF2a 的基因与英国分离株 PLRV-S、荷兰分离株 PLRV-N、澳大利亚分离株 PLRV-A、加拿大分离株 PLRV-C 的 ORF2a 基因的序列同源性分别为 98.96%、98.70%、94.79%、97.5%,其中与 PLRV-S 的同源性最高,与 PLRV-A 的同源性最低。ORF2a 5'端的同源性相对比 3'端的同源性要高,特别是在 ORF2a 与 ORF2b 重叠区附近与移码有关的滑动序列的上游非同源性最高。在相当于 PLRV-S 的 1770 核苷酸左右有一个 7 核苷酸组成的与移码有关的滑动序列(图 1 中,下面划横线)UUUAAA,其下游有一个茎环结构(图 1 中下面划有浪线)。其上游相距 5 个核苷酸处,有一个 PLRV RNA 特有的连续排列的 3 次重复序列(见图 1 下面划点线)和 Mayo 等(1989)报道是一致的。每段由 27 个核苷酸组成,为 PLRV 各分离株所特有。我们推测这种 3 次重复序列可以形成复杂的结构,构成较紧密连续折叠互补双链结构和发夹结构(图

2)。其作用可能和 -1 移码有关。

2.3 ORF2a 编码的 70kD 的蛋白的氨基酸序列

根据核苷酸序列分析推测 PLRV ORF2a 编码的 70kD 的蛋白的氨基酸序列见图 3。分析结果表明中国株 PLRV-Ch 编码的氨基酸与英国分离株 PLRV-S、荷兰分离株 PLRV-N、澳大利亚分离株 PLRV-A、加拿大分离株 PLRV-C 的 ORF2a 基因编码的氨基酸的序列同源性分别为 97.97%、97.97%、89.69%、95.94%,其中与 PLRV-S、PLRV-N 的同源性最高,与 PLRV-A 的同源性最低。在该蛋白的 C 端有一个蛋白酶的特征序列 H(X25)D/E(X70-80)R/K(XGXSG)(图 3 中划有横线),推测此结构具有剪切 V_{pg} 前体的功能;除此之外,还有一个螺旋酶的特征序列(motif IV)NYVFESTA(图 3 中划有浪线),其功能尚不清楚。

2.4 PLRV-Ch 在 ORF2a 与 ORF2b 重叠区经移码后编码的转码蛋白的氨基酸序列

根据核苷酸序列分析推测 PLRV ORF2a 在 ORF2a 与 ORF2b 重叠区经移码后编码的转码蛋白的氨基酸序列见图 4 发现该段序列有一个富含精氨酸、赖氨酸、组氨酸的区域 KRQLRHPRRRYK(图 4 中划横线),与 PLRV 的 17kD 的核酸结合蛋白的碱性区很类似,可能与复制时结合核酸有关。

3 讨论

ORF2a 的二级结构比较复杂有各种发夹结构和互补双链区,可能是聚合酶作用的调控部位。

与移码有关的滑动序列 UUUAAA,其下游有一个茎环结构,其上游相距 5 个核苷酸处,有一个 PLRV RNA 特有的连续排列的 3 次重复序列,每段由 27 个核苷酸组成,为 PLRV 各分离株所特有。我们推测这种 3 次重复序列可以形成复杂的结构,构成较紧密连续折叠互补双链结构和发夹结构,其作用可能和 -1 移码有关。色氨酸操纵子的衰减子中也有类似的多重链内互补结构^[15],它对色氨酸合成酶的基因的转录起衰减作用。在 HBV(乙肝病毒)的 adr NC-1 DNA 中也有类似的多重链内互补结构^[16],包括一个以上的小肽框架,与衰减子行使衰减作用调控的结构特点非常相似,是调控的重要部位。所以推测此连续折叠互补双链结构和发夹结构,可能与基因的表达调控有关。ORF2a 编码 70kD 的蛋白,具有蛋白酶的特征序列和螺旋酶的特征序列;除此之外,在 ORF2a 与 ORF2b 重叠区经移码后编码的转码蛋白具有富含精氨酸、赖氨酸、组氨酸的区域 KRQLRHPRRRYK,这些特征性的序列具有特殊的功能。因此对 ORF2a 基因结构的研究与分析,对于我们进一步研究 ORF2a 基因的表达调控以及所编码的蛋白的功能具有重要的作用。ORF2a 编码的蛋白,在复制酶的复制中起一定的作用^[17],是必须的,若以 ORF2a 5'的反义 RNA 转化马铃薯,封闭 ORF2a 基因,ORF2a 不能正常表达,可干扰聚合酶的正常作用,从而获得抗 PLRV 的马铃薯植株。所以 ORF2a 基因的获得,为今后研究马铃薯抗病毒工程开辟了新的途径。

PLRV-Ch	AGGAACAGAU	UTACCCGAUA	DGCGGCUCU	UUCUUADAU	UCGCCUUUG	PLRV-Ch	CCAUAGAUGG	AGCCCADAC	CAGUUCGUGU	CGGUUUADG	CAACACCGGA
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1307-1356
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1203-1252
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1204-1253
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1204-1253
PLRV-Ch	CCCAACUGCA	AAAGAGGCCG	GAUUUCAACA	UCGCGUCUC	AACUUCGAG	PLRV-Ch	CCCGGAAUAT	CCGGAACAGG	GUUUGGUCU	UCAGAAAUUC	UCGUUGUGU
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1357-1406
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1253-1302
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1254-1303
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1254-1303
PLRV-Ch	GCACUCAC	UAUGAGUCGC	UUGAGUGGGG	AUACUCUCG	GGCACCCACC	PLRV-Ch	GCUTAAAGG	UUCACUCUG	AGAGGAGGG	UAACCAAAU	GUUUGUCUG
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1407-1456
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1303-1352
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1304-1353
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1304-1353
PLRV-Ch	CCGCGADACA	AADCGDGGG	CCUACCCAGC	UCACCAAAU	UGAGAGCCA	PLRV-Ch	UUUUAACCCU	GADCCAGGA	ADACUUCUC	CAAAUUAU	GUUUGAGUG
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1457-1506
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1353-1402
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1354-1403
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1354-1403
PLRV-Ch	ACCAACCGCG	CGCGUACAG	AUCGAGCUGA	CUAGAGUUA	GUUCAAGUC	PLRV-Ch	ACCAGCGUAA	AGGCGCGUG	CUUCGCGAG	GAAGCGUGA	AAAGCUGAA
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1507-1556
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1403-1452
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1404-1453
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1404-1453
PLRV-Ch	UUUAUCCAA	AAAGCGGCG	GAUUUGCAAA	CGGUUGGGG	CATGACAGG	PLRV-Ch	CGGGGAAGCA	UCCGAAGCCG	UCAAGAGCCU	UGCCAGAUU	AAUACACUA
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1557-1606
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1453-1502
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1454-1503
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1454-1503
PLRV-Ch	AGCAUUGU	CAGAAAGCU	AUUUGCCUC	UGGAACCCG	DGAAAGAAU	PLRV-Ch	CCGCGAAGAA	CCGGGCGGAT	GAUUUGAGU	CCGAGAGGA	UUAAGGUCU
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1607-1756
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1503-1652
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1504-1653
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1504-1653
PLRV-Ch	ADCCUCAAA	GGCGCCUCG	UGACCUUAG	GGCAUUUAC	AGCAUUUGU	PLRV-Ch	GAGAGAGAG	CCGCAACAA	CGCGCCGCA	GAGAAAGCC	UCUAAACAA
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1757-1806
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1653-1702
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1654-1703
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1654-1703
PLRV-Ch	UCGUCUCDA	UAAGAGCGU	GCAAGGUGA	UCACUUUGU	CCUCGGAGU	PLRV-Ch	CUACAGAGAG	AAGACGUCU	CAUCAAUCU	AGCAGAGAA	ACTGUCGCA
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1807-1856
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1703-1752
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1704-1753
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1704-1753
PLRV-Ch	UUCAGAUAG	AAAGCUUAG	CUUAAUUGU	CUUGGUGUA	UAACGAGUU	PLRV-Ch	CAAAACAGCC	UUUUAAGGG	CAAGCGCAC	CGUCGCGCA	AACAAGCGC
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1857-1906
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1653-1702
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1654-1703
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1654-1703
PLRV-Ch	GAUUACAG	GGCGCGUAA	GUUUUCAGA	GAACUUACC	GUUUUCUGU	PLRV-Ch	AACUCGACA	UCCCGGAGC	CGCUACAGG	GCACCCCAA	UGAGAAAUU
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1807-1856
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1703-1752
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1704-1753
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1704-1753
PLRV-Ch	UUUAUCUC	UCUGAAGU	AUUUGAGGG	CAGCUUCUC	CAAAAGAAU	PLRV-Ch	GGUGAACAG	AUCAACAG	CUAUGGGGG	GAGAAUAAU	CUUCUGAGA
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1857-1906
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1753-1802
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1754-1803
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1754-1803
PLRV-Ch	UACAGAAG	AGAAGCUCU	GGAGGAGAC	AAAGGUUUU	CGGUCGACA	PLRV-Ch	UAGAGAGAA	GAUAGGAGC	AGGUGGUCU	AGAAAGCCU	CGAGAGCCU
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1907-1956
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1803-1852
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1804-1853
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1804-1853
PLRV-Ch	AAAACCGCA	AAAGCUCUC	UAAUAGAAU	ACAACAGAA	AACGCGAGC	PLRV-Ch	AAACAAAGA	AGCGCGAAG	CGUGGAGGG	AAGAACAAG	AAAACAGUC
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	1957-2006
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1853-1902
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1854-1903
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1854-1903
PLRV-Ch	AUCUCGGUA	CGGACAGCG	AUCUCUCUG	ACAGUAGGA	GAACCCUUG	PLRV-Ch	ACCUCUACU	UCAACGCAU	CUACAAGUG	GGCGCCCAAG	AAAGGCGUG
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	2007-2056
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1903-1952
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1904-1953
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1904-1953
PLRV-Ch	GGAGAGCUG	AACACGUCU	AGAAGGCGU	UUCGCAAGU	CGUUGAAAC	PLRV-Ch	CCCCCAGCG	UUCAGGAAG	GGCGGACAU	CCCGGUCUC	UACCAACCCU
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	2057-2106
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	1953-2022
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	1954-2003
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	1954-2003
PLRV-Ch	UGAAACAGG	AUCCGAGUG	CGACUUUCU	UCCCAUUUC	AAAAGGCCC	PLRV-Ch	GCACAGAGG	CGAAACCCG	UGGGGCAAA	AACUCUCCA	AGUUCACCC
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	2107-2156
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	2003-2052
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	2004-2053
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	2004-2053
PLRV-Ch	GUAAUGAAU	CUCCAGACU	GUAGGUCAC	CCACCGGGA	AGGUUACUA	PLRV-Ch	GAGCGGCGG	AGAAACAGC	AGGAUUCGC	UGGCGAAAG	CCGGAUCUA
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	2157-2206
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	2053-2102
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	2054-2103
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	2054-2103
PLRV-Ch	UCAGUCAAG	GAGCUCADU	CAUACAGCC	GACAAAACG	GCAAAAGUCC	PLRV-Ch	AGCUGAACUC	CAAAGCCUA	-----	-----	-----
PLRV-S	-----	-----	-----	-----	-----	PLRV-S	-----	-----	-----	-----	2207-2226
PLRV-N	-----	-----	-----	-----	-----	PLRV-N	-----	-----	-----	-----	2103-2122
PLRV-A	-----	-----	-----	-----	-----	PLRV-A	-----	-----	-----	-----	2104-2123
PLRV-C	-----	-----	-----	-----	-----	PLRV-C	-----	-----	-----	-----	2104-2123

图 1 PLRV-Ch ORF2a 基因核苷酸序列及其与国外 4 种 PLRV 分离株同源性比较

Fig.1 Nucleotide sequence of PLRV-Ch ORF2a gene and the homology comparison with foreign isolates , PLRV-S , PLRV-N , PLRV-A , PLRV-C

- FEBS Letters*, 1989 **245** 51 ~ 56
- [4] ZHANG H I(张鹤龄). Development of potato leafroll virus genome. *VIROLOGY SINICA(中国病毒学报)*, 1996 **11**(1):1 ~ 8
- [5] Prufer D, Tacke E, Schmitz J *et al.* Ribosomal frameshifting in plants: a novel signal directs the -1 frameshifting in the synthesis of the putative viral replicase of potato leafroll luteoviruses. *EMBO J*, 1992 **11**(3):1111 ~ 1117
- [6] Kujawa A B, Drugron G, Hulanicka D *et al.* Structural requirement for efficient translational frameshifting in the synthesis of the putative viral RNA-dependent RNAPolymerase of potato leafroll virus. *Nucleic Acids Res*, 1993 **21**(9):2165 ~ 2171
- [7] Gorbalenya A E, Donchenko A P, Blinov V M *et al.* Cysteine protease of positive strand RNA viruses and chymotrypsin-like serine proteases. *FEBS LETT*, 1989 **243**:103
- [8] Tacke E, Prufer D, Schmitz J, Rohde W. The potato leafroll luteovirus 17kD protein is a single-stranded nucleic acid bind protein. *J Gen Virol*, 1991 **72**:2035 ~ 2038
- [9] Golemboski D B, Lomonosoff G P, Zaitlin M. Plants transformed with a tobacco mosaic nonstructural gene sequence are resistant to the virus. *Proc Natl Acad Sci USA*, 1990 **87**:6311 ~ 6315
- [10] HASI Agula(哈斯阿古拉), SHI Y(施一), ZHANG H I(张鹤龄). The cDNA synthesis, molecular cloning and nucleotide sequence of potato leafroll virus coat protein gene. Proceedings of Asia-Pacific Conference on *Agriculture Biotechnology(农业生物技术)*, Beijing, China, 1992, pp. 153 ~ 156
- [11] ZHANG R X(张荣信), HASI Agula(哈斯阿古拉), ZHANG H L(张鹤龄). Cloning and sequence of 56kD protein gene and 3' non-coding region of potato leafroll virus. *Virology(病毒学报)*, 1997 **13**(3):247 ~ 25
- [12] DONG J L(董江丽), HASI Agula(哈斯阿古拉), ZHANG H L(张鹤龄). Cloning and sequence of gene interval region of potato leafroll virus. *VIROLOGY SINICA(中国病毒学报)*, 1996 **11**(2):144 ~ 148
- [13] LIANG C(梁成罡), HASI Agula(哈斯阿古拉), ZHANG H L(张鹤龄). Cloning and sequence of replicase gene 3' region of potato leafroll virus. *Virology(病毒学报)*, 1997 **13**(3):278 ~ 282
- [14] MENG Q(孟清), ZHANG H I(张鹤龄). Purification of potato leafroll virus. *J Virology(病毒学报)*, 1987 **2**(2):151 ~ 155
- [15] ZHU Y X(朱玉贤), LI Y(李毅). Modern Microbiology. First printing, Beijing: High education publish agency, 1997, p142
- [16] SHEN Q F(沈鲈), FANG D F(方德福) *et al.* Expressing and regulation of eucaryon gene. First printing, Beijing: High education publish agency, 1996, p196
- [17] Ahlquist P, Strauss E G, Rice C M *et al.* Sindbis virus protein nsp1 and nsp2 contain homology to nonstructural protein from several RNA plant viruses. *J Virology*, 1985 **53**:536 ~ 542

Structure Characteristics of ORF2a Gene of Potato Leafroll Virus Chinese Isolate

ZHAO Guo-Fen² ZHANG He-Ling^{1*} HASI Agula¹

¹(Department of Biology, Inner Mongolia University Hohhot 010021, China)

²(Department of Biological Engineering, Inner Mongolia Agriculture University, Hohhot 010018, China)

Abstract According to the genomic sequence of foreign four PLRV isolates, three pairs of specific primer were designed and synthesized. The cDNA of the ORF2a gene of PLRV-Ch was synthesized by reverse transcription and followed by Polymerase Chain Reaction amplification. The synthesized 3' and 5' cDNA fragment of the PLRV-Ch ORF2a gene were inserted into pUC19 and cloned in *E. coli* JM109 and were sequenced respectively. The middle cDNA fragment were directly sequenced. The homology of nucleotide sequence of PLRV-Ch compared with PLRV-S(Scotland, UK), PLRV-N(Netherlands), PLRV-A(Australia) and PLRV-C(Canada) were 98.96%, 98.70%, 94.79%, 97.5%, the homology of putative amino acid sequence are 97.97%, 97.97%, 89.69%, 95.94%. In 3' region of ORF2a gene a slippery sequence for -1 frameshift and its downstream "stem-loop" or "pseudoknot" and upstream nucleotide sequence repeats were found. Authors suggested that the nucleotide repeat sequences characteristic for PLRV could form a tight successively folded complementary double stranded regions and hairpins. This structure possibly has something to do with -1 frameshift. The amino acid sequence of C terminus region of 70kD protein translated by motif IV has a protease characteristic motif and a helicase motif IV. The amino acid sequence of polypeptide translated by ORF2a gene undergoing frameshift has a single-stranded nucleic acid binding protein-like characteristic motif.

Key words Potato leafroll virus, gene cloning, nucleotide sequence, ORF2a

Received: 05-31-2002

This work was supported by Grant from National Natural Sciences(No.9270453).

* Corresponding author. Tel 86-471-4992693; Fax 86-471-4992442; E-mail: hlzhang@nmg2.imu.edu.cn