

# 番木瓜环斑病毒复制酶(亚基)基因的克隆、序列分析 及其植物表达载体的构建

刘俊君 彭学贤 莽克强

(中国科学院微生物研究所植物生物工程开放实验室, 北京 100080)

近二年报道,在 TMV<sup>(1)</sup>、PVX<sup>(2)</sup>、PEBV<sup>(3)</sup>和 CMV<sup>(4)</sup>等病毒中,利用病毒编码的复制酶(亚基)基因或相关 cDNA 片段转化烟草,工程植株获得绝对或极高的病毒抗性。大量研究认为:马铃薯 Y 病毒组的核内含体大分子量蛋白(NiB)是依赖于 RNA 的 RNA 复制酶(或核心亚基),NiB 与上述病毒的复制酶有广泛的同源保守区<sup>(5)</sup>。因此,NiB 基因的克隆,不但在植物抗病基因工程方面,而且在马铃薯 Y 病毒组基因组的复制研究方面,均有重要的意义。本文以马铃薯 Y 病毒组的重要成员番木瓜环斑病毒的华南强株系(PRSV-SM)为材料,克隆了 PRSV 的 NiB 基因,并完成其全序列测定和植物表达载体的构建,为探索马铃薯 Y 病毒组复制酶可能介导的抗病性打下了基础。

## 1 材料和方法

### 1.1 材料

病毒株系由中国科学院华南植物研究所朱西儒先生惠赠。大肠杆菌菌株和质粒等均由本实验室提供。cDNA 合成试剂购自 Boehringer 公司,T7 DNA 聚合酶序列分析试剂盒购自 Pharmacia 公司,各种限制酶购自华美公司, $\alpha$ -<sup>32</sup>P-dATP 等同位素试剂为 NEN 产品,寡聚脱氧核苷酸由本所技术室合成。

### 1.2 方法

1.2.1 病毒提取以及病毒 RNA 纯化:参照文献 [6]。

1.2.2 cDNA 合成及其克隆:以 PRSV-RNA 为模板,人工合成的寡聚核苷酸作为第一条链合成的引物:5'TTACTCATTCAAACCAGCAT3'。引物是根据 PRSV-SM 外壳蛋白基因 5'端区序列<sup>(7)</sup>设计的,同时添加了终止密码。合成的 ds-cDNA 经电泳分离,收回 1.0—3.0kb 的片段,平头连入 pBluescript SK 的 EcoR V 位点,转化大肠杆菌。在含有 IPTG 和 X-gal 的氨苄青霉素 LB 平板上,挑选白色重组克隆。

1.2.3 NiB 基因克隆的筛选:对已报道 PRSV-W 和 PRSV-HA5-1 株系 NiB 基因 3'片段的序列<sup>(8)</sup>进行了分析,其 3'端区有保守的 Hind III 位点。据此对重组克隆进行 Hind III 位点测试,挑出与国外株系有相同 Hind III 位点的重组克隆,进行两端序列测定来确证其是否为 NiB 基因片段。在此基础上对它们进行插入大小和酶谱分析,挑出最大的重组克隆。

1.2.4 DNA 序列分析:根据酶谱进行亚克隆,按常规双脱氧链终止法进行序列测定。

1.2.5 NiB 基因 5'端区域的 PCR 法修饰:根据本文所测序列和 PRSV-HA<sup>(9)</sup>序列,设计寡聚核苷酸引物,进行了 30 个循环的 PCR。94℃变性 1 分钟,50℃复性 1 分钟,72℃延伸 1 分钟。PCR 产物经酚抽提后,乙醇沉淀回收。

本文系国家“八·五”攻关项目资助课题,曾在 1992 年 10 月第四届全国病毒生化学术讨论会上报告。  
本文于 1993 年 7 月 15 日收到。

## 2 结果和讨论

### 2.1 cDNA 的克隆和 Nib 基因组克隆的鉴定

我们开始克隆 Nib 基因时,国际上尚无 PRSV 完整 Nib 基因克隆和序列报道,因此本实验采用了 ds-cDNA 克隆到载体,构建质粒库的方法。本实验共挑选到 76 个白色重组克隆。利用载体的 BamH I 和 Sal I 双酶切分析,插入片段长度在预期大小范围内。对各重组克隆进行 Hind III 位点测试,其中 4 个克隆有保守的 Hind III 位点,位于插入片段一侧的约 300bp 处。这与 PRSV 国外株系已发表的部分 Nib 基因片段内 Hind III 的位置相当。其中,第 75 号重组克隆(pRS75)插入片段最长,约 1.5kb,同马铃薯 Y 病毒组 Nib 基因的一般大小相近,部分序列测定结果表明它为 Nib 基因的 cDNA 片段。对 pRS75 进行酶谱分析,其结构如图 1 所示。

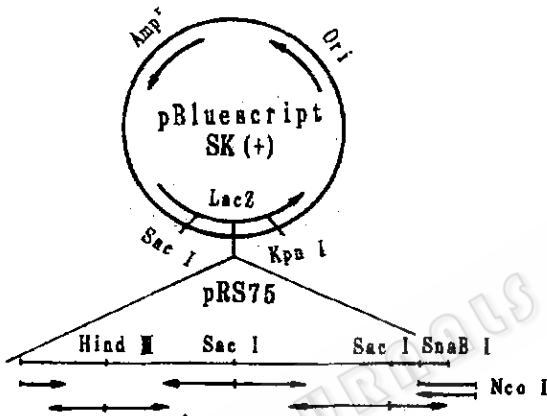


图 1 pRS75 重组质粒结构图  
(箭头表示测序策略)

### 2.2 Nib 基因的 cDNA 序列

利用 pRS75 插入片段内的一个 Hind III 和两个 Sac I 位点,构建 4 个亚克隆,完成其全序列分析。结果表明该克隆含有 Nib 基因的 1490bp,与马铃薯 Y 病毒组其他成员比较,尚缺 5'端的少量序列。为了补齐这部分序列,并对 Nib 基因 5'端进行了必要的修饰,增添起始密码和 Kozak 序列。根据最近发表的 PRSV-HA<sup>69</sup> Nib 基因序列设计 PCR 的 5'端引物: ACCATGGGTGGTGGTTCGTTGGCTCT3';

根据 pRS75 的序列,在其 5'端的 SnaB I 位点区设计 PCR 的 3'端引物: 5'GTCCTT-TACGTATGCTTC3'。从 cDNA 中扩增出长约 200bp 的 Nib 基因 5'端区域,克隆到 pBlue-script SK 的 EcoR V 位点内,得重组克隆 pRB

+1,测定其序列。这样获得了 PRSV-SM 株系 Nib 基因的全序列(如图 2)。其全长 1611bp,编码 537 个氨基酸。Nib 基因和其上游的 Nla 基因以及其下游的外壳蛋白基因之间连续通读。这与马铃薯 Y 病毒组基因组的一般结构特点一致。

### 2.3 Nib 基因植物表达载体的构建

将 pRS75 中插入片段在 pGEM 5Zf (+) 中调头之后,通过 SnaB I 位点完成 NIB 全基因的拼接。利用载体的 BamH I 和 Sal I 位点,将该基因插入 pBI121 的衍生载体 pBin438 中,得到 PRSV-Nib 基因的高等植物表达载体 pRB1.6 (见图 3)。

```

----->N1b
S  G  G  R  W  L  P  D  K  L  H  G  N  L  K  G  V  S  S  A  S  S  N  L
AGT GGT GGT CGT TGG CTC TTT GAC AAA TTA CAC GGC AAT CTC AAG GGA GTT AGC TCA GCT TCC AGC AAT TTA
V  T  K  H  V  V  K  G  T  C  P  L  P  R  N  Y  L  E  C  D  E  E  A  K
GTC ACG AAG CAT GTT GTG AAA GGC ACT TGT CCT CTC TTC AGG AAT TAT CTT GAG TGT GAT (AAA GAA GCT AAG
T  F  Y  S  P  L  M  G  H  Y  M  K  S  V  L  S  K  E  A  Y  V  K  D  L
ACT TTC TTT AGT CCG CTG ATG GGA CAC TAC ATG AAG AGC GTT CTA AGC AAG GAA GCA TAC GTA AAG GAC TTA
M  K  Y  S  S  D  I  I  V  G  E  V  D  H  D  V  F  E  E  S  V  T  Q  V
ATG AAG TAT TCA AGT GAC ATA ATT GTT GGT GAG GTT GAT CAT GAT GTG TTT GAG GAA AGC GTT ACA CAA GTT
V  E  L  L  N  D  H  E  C  P  E  L  E  Y  I  T  D  S  E  V  I  I  Q  A
GTC GAG TTG CTA AAC GAC CAC GAG TGT CCT GAG CTC GAA TAC ATC ACA GAT AGC GAA GTC ATC ATA CAA GCT
L  N  M  D  A  A  V  G  A  L  Y  T  G  K  K  R  K  Y  F  E  G  S  T  V
CTG AAT ATG GAT GCT GCT GTC GGA GCT TTG TAC ACA GGA AAG AAA AGG AAA TAT TTC GAA GGA TCA ACA GTT
E  H  R  Q  A  L  V  R  K  S  C  E  R  L  Y  E  G  R  M  G  V  W  N  G
GAG CAC AGA CAA GCC TTA GTT CGA AAG AGC TGT GAG CGT TTG TAT GAA GGG CGA ATG GGT GTT TGG AAT GGA
S  L  K  A  E  L  R  P  T  E  K  V  L  A  K  K  T  R  S  F  T  A  A  P
TCA CTA AAG GCT GAG CTG CGA CCT ACA GAG AAG GTC TTG GCC AAA AAG ACA AGA TCT TTC ACT GCA GCG CCC
L  D  T  L  L  G  A  K  V  C  V  D  D  F  N  N  W  F  Y  S  K  N  M  D
CTT GAT ACT TTG TTA GGA GCC AAA GTT TGT GTT GAT GAT TTT AAT AAT TGG TTC TAC AGC AAA AAT ATG GAT
C  P  W  T  V  G  M  T  K  F  Y  K  G  W  D  E  F  L  R  K  F  P  D  G
TGT CCT TGG ACT GTT GGT ATG ACA AAA TTT TAT AAA GGC TGG GAC GAA TTT TTG AGA AAA TTT CCT GAT GGT
W  V  Y  C  D  A  D  G  S  Q  F  D  S  S  L  T  P  Y  L  L  D  A  V  L
TGG GTA TAT TGC GAT GCA GAT GGA TCT CAG TTT GAT AGC TCT TTG ACA CCG TAC TTG TTA GAT GCT GTG TTA
S  I  R  L  W  A  M  E  D  W  D  I  G  A  Q  M  L  K  N  L  Y  G  E  I
TCA ATT CGA CTG TGG GCT ATG GAA GAT TGG GAC ATT GGA CCT CAG ATG CTT AAG AAC CTA TAT GGA GAG ATC
T  Y  T  P  I  L  T  P  D  G  T  I  V  K  K  F  K  G  N  N  S  G  Q  P
ACA TAC ACA CCT ATT TTG ACA CCT GAT GGG ACA ATT GTC AAG AAA TTC AAA GGC AAT AAT AGT GGA CAA CCT
S  T  V  V  D  N  T  L  M  V  L  I  T  M  Y  Y  A  L  R  K  A  G  Y  D
TCG ACA GTT GTT GAT AAC ACA CTG ATG GTT TTG ATC ACA ATG TAC TAC GCG TTA CCG AAG GCT GGT TAC GAC
A  K  A  Q  E  D  M  C  V  F  Y  I  N  G  D  D  L  C  I  A  I  H  P  D
GCG AAA GCT CAG GAG GAT ATG TGC GTG TTT TAT ATC AAT GGT GAT GAT CTC TGT ATT GCC ATT CAC CCA GAT
H  E  H  V  L  D  S  F  S  S  S  F  A  E  L  G  L  K  Y  D  F  T  Q  R
CAT GAA CAT GTT CTT GAC TCA TTC TCT AGT TCA TTT GCT CAG CTT GGA CTT AAA TAT GAT TTC ACT CAA AGA
H  R  N  K  Q  D  L  W  F  M  S  H  R  G  I  L  I  D  D  V  Y  I  P  K
CAC CGA AAT AAA CAA GAT TTG TGG TTT ATG TCG CAT CGA GGT ATT CTG ATC GAT GAT GTT TAT ATT CCC AAA
L  E  P  E  R  I  V  A  I  L  E  W  D  K  S  K  L  P  E  H  R  I  E  A
CTT GAA CCT GAA AGA ATT GTG GCT ATC CTT GAG TGG GAT AAA TCT AAT GTT CCA GAA CAC CGA TTA GAA GCT
I  T  A  A  M  I  E  S  W  G  Y  G  E  L  T  H  Q  I  R  R  F  Y  Q  W
ATC ACA GCA GCA ATG ATA GAG TCA TGG GGA TAC GGG GAG TTA ACG CAC CAG ATT CGT AGG TTT TAT CAA TGG
V  L  E  Q  A  P  F  N  E  L  A  K  Q  G  R  A  P  Y  V  S  E  V  G  I
GTG CTT GAG CAG GCT CCA TTC AAT CAG TTA GCA AAG CAG GGA AGG GCT CCT TAT GTG TCC GAG GTT GGA TTG
R  R  L  Y  T  S  E  R  G  S  M  D  E  L  E  A  Y  I  D  K  Y  F  E  R
AGG CCG TTA TAC ACT AGT GAG CGT GCG TCT ATG GAT CAA TTA GAA GCC TAC ATA GAT AAA AAT TTT GAG CGA
E  R  G  D  S  P  E  L  L  V  Y  H  E  S  R  S  T  G  N  C  Q  F  V  C
GAA AGG GGA GAT TCA CCT GAG CTA CTG GTG TAT CAT GAA TCG ACA AGC ACT GGC AAT TGT CAG TTT GTC TGC
G  S  D  T  H  V  F  H  Q  S  K  N  E  A  V  D  A  G  L  N  E
GGC AGT GAC ACA CAT CTG TTT CAC CAG/TCC AAA AAT GAA GCC GTG GAT GCT GTT TTG AAT GAG
N1b--|--CP
    
```

图2 PRSV-SM N1b 基因核苷酸序列及其推导出的氨基酸序列  
“/”表示 N1b 和 CP 之间的分割点

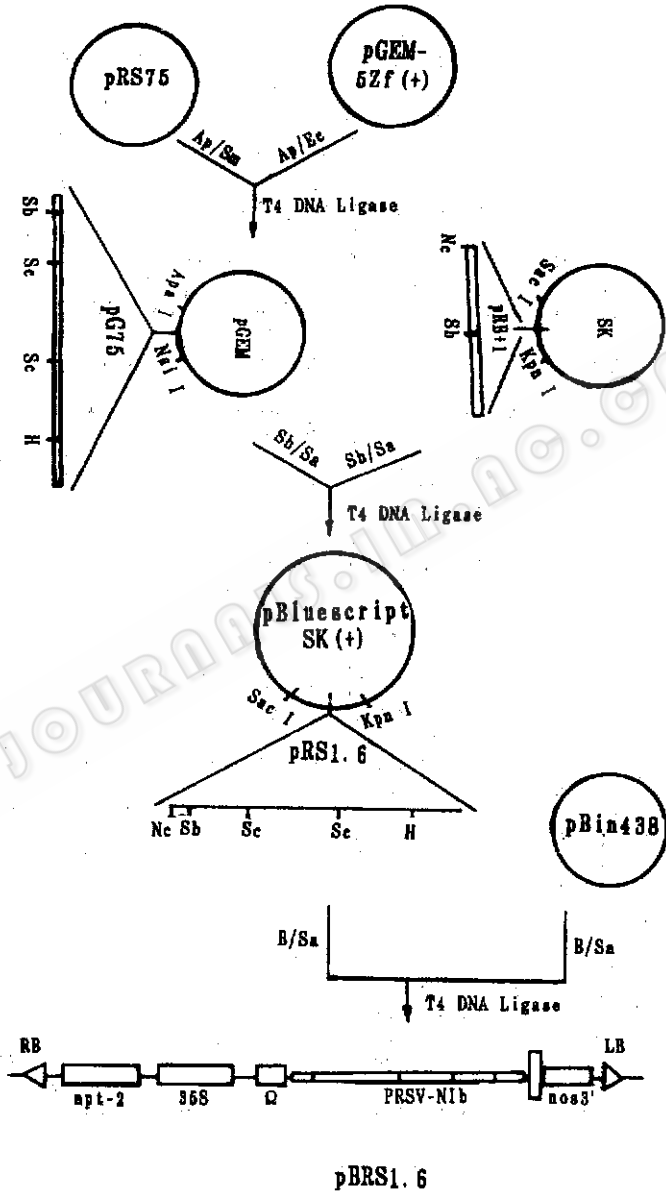


图3 PRSV-N1b 基因高等植物表达载体的构建

内切酶简写如下: Ap. Apa I; Sm. Sma I;

Ec. EcoR V; Nc. Nco I; Sb. SnaB I; Sc. Sac I; Sa. Sal I; H. Hind III

## 参 考 文 献

- [1] Daniel B *et al.* Proc Natl Acad Sci USA, 1990, 87: 6311—6315.
- [2] Carl J B *et al.* The Plant Cell, 1992, 4: 735—744.
- [3] Stuart A *et al.* Proc Natl Acad Sci USA, 1992, 89: 5829—5833.
- [4] Anderson J M *et al.* Proc Natl Acad Sci USA, 1992, 89: 8759—8763.
- [5] Eugene V K. Journal of General Virology, 1991, 72: 2197—2206.
- [6] 刘俊君等. 病毒学报, 1993, 6(1): 67—81.
- [7] 刘俊君等. 微生物学通报, 1991, 18(6): 350—351.
- [8] Quemada H *et al.* Journal of General Virology, 1990, 71: 203—210.
- [9] Yeh S-D *et al.* Journal of General Virology, 1992, 73: 2531—2541.

# Study on Replicase (Subunit) Gene of Papaya Ringspot Virus: Cloning, Sequencing and Construction of Higher Plant Expression Vector

Liu Junjun Peng Xuexian Mang Keqiang

(Laboratory of Plant Biotechnology, Institute of Microbiology, Academia Sinica, Beijing 100080)

**Abstract** The ds-cDNA was synthesized using genomic RNA of papaya ringspot virus (PRSV) as template. The blunt-ended ds-cDNA was cloned into the EcoRV site of vector pBluescript SK. From the recombinants, the N1b gene of PRSV was obtained and its complete sequence was determined. After the N1b gene was modified through PCR, a binary vector of PRSV-N1b gene has been constructed for its expression in higher plant.

**Key words** Papaya ringspot virus, N1b gene, higher plant expression vector