

黄瓜花叶病毒香蕉株系(CMV-Xb)RNA3 cDNA的克隆和序列分析

王海河, 谢联辉**, 林奇英

(福建农业大学植物病毒研究所; 福建省植物病毒学重点实验室, 福建福州 350002)

Construction of cDNA Clone and Nucleotide Sequence of RNA3 of
Cucumber Mosaic Virus Isolate Xb from Banana

WANG Hai-he, XIE Lian-hui**, LIN Qi-ying

(Institute of plant virology, Fujian Agricultural University, Key Lab of Plant Virology of Fujian Province, Fuzhou 350002, China)

Abstract: The nucleotide sequence of RNA3 of a cucumber mosaic virus strain from banana of subgroup II (CMV-Xb) was cloned and analyzed at nucleotide and protein level after RT-PCR amplification with two pairs of primers. The results showed that CMV-Xb is composed of 2205 nucleotides (nt) with two open reading frames. The one near 5' termini (97 ~ 936nt) encodes a 3a protein with 279 amino acids (aa), the other near 3' termini (1225 ~ 1871nt) codes for coat protein with 218aa. The 5' non-coding region (NR), intercistron region (IR) and 3' NR are 96nt, 288nt and 324nt, respectively. Comparison at the nucleotide and putative amino acid level with other strains in subgroup II showed that the nucleotide sequence of Xb is very conservative in the encoding regions as well as non-coding regions. The evolution of RNA3s of subgroup II members of CMV is consecutive.

Key words: Cucumber mosaic virus strain from banana; RNA3; Full-length cDNA; Nucleotide sequence

摘要: 通过 RT-PCR 方法, 设计两对引物, 克隆了黄瓜花叶病毒香蕉株系 (CMV-Xb) RNA3, 并进行了核苷酸和蛋白质水平上的分析。结果表明 Xb 株系 RNA3 全长 2205nt, 具有两个蛋白编码阅读框架 (ORF), 其中 5' 端 (97 ~ 936nt) 编码一个 279aa 的 3a 蛋白; 3' 端 (1225 ~ 1871nt) 编码一个 218aa 的 CP 蛋白。5' 非编码区域为 96nt; 基因间隔区 (IR) 长 288nt; 3' NR 含有 324nt。通过与亚组 II 其它株系 RNA3 核苷酸和所编码产物推导的氨基酸序列分析发现, 亚组 II 株系无论在编码区还是非编码区的核苷酸同源性都相对较高; 亚组 II 株系在进化过程中具有连续性。

关键词: 黄瓜花叶病毒香蕉株系; RNA3; 全长 cDNA 克隆; 序列分析

中图分类号: S432.4 文献标识码: A 文章编号: 1003-5125(2001)03-0252-05

黄瓜花叶病毒 (CMV) 是黄瓜花叶病毒属的典型成员, 为正链三基因组的分体病毒^[1]。研究表明, CMV RNA1, 2 分别编码与复制酶有关的 1a 和 2a 蛋白^[2], 此外 RNA2 还编码一个与病毒基因表达有关

的 2b 蛋白^[3]; RNA3 编码 3a 运动蛋白和外壳蛋白, 但外壳蛋白一般由 RNA3 3' 端亚基因组化产生的 RNA4^[4] 所编码表达。CMV 的寄主范围极其广泛, 是农作物及经济观赏植物的主要病毒之一, 为了从分

收稿日期: 2000-08-14, 修回日期: 2001-01-10

作者简介: 王海河 (1970 -), 男, 陕西永寿县籍, 博士, 主要从事植物分子病毒学研究。

** 通讯作者。Correspondence author.

本文报道的核苷酸序列已被贮存于 GenBank 数据库, 接受号 AF268598

子水平上对 CMV 的致病机制进行深入研究,我们克隆并分析了从香蕉上分离的经血清学、生物学和 CP 基因特征证实为 CMV 亚组 II 的株系 Xb^[6]的 RNA3 全序列特征,现报道如下。

1 材料与方法

1.1 毒源来源

CMV-Xb 株系由厦门华侨亚热带引种植物园徐平东博士惠赠,并保存在本实验室。

1.2 病毒及其 RNA 的提取

用 CMV-Xb 接种心叶烟 (*N. glauca*), 14d 后采集接种叶以上出现症状的叶片, -70℃ 保存。病毒提纯参照 Mossop 等^[7]的方法。病毒 RNA 的提取依照 Peden^[1]的方法。用 DEPC 处理过的灭菌重蒸水溶解病毒 RNA, -70℃ 保存。

1.3 RNA3 cDNA 第一链的合成

由于 RNA3 比较长,因此将 RNA3 分为两个片段进行克隆。根据已发表的 CMV 不同株系 RNA3 的序列分析表明,CMV-RNA3 的 5' 和 3' 端均很保守,因此根据株系 Q 的序列合成 RNA3 3' 端和一个中间 3' 端引物;3' 端引物 a: 5' CCGGATCCTG-GTCTCCTTATGGAG 3' 和中间引物 E1: 5' CTA CTAG-CATTGGGAG 3' (1251 ~ 1236nt)。反应体系中含 RNA3 2.5μL (0.4 ~ 0.5μg/μL); 3' 端引物 a 和中间引物 E1 1μL (10pmol/μL); 重蒸水 6.5μL, 90℃ 水浴 10min, 立即冰上 5min, 然后加入 5X 反转录缓冲液 5μL; RNA 酶抑制 1μL (40U/μL); MuMLV 反转录酶 1μL (20U/μL); dNTPs 2.5μL; DDW 5.5μL。37℃ 保温 1h, 然后 95℃ 5min 灭活反转录酶。

1.4 cDNA 第二链的合成和 PCR 扩增

反应体系包括 cDNA 5μL, dNTPs 1μL; 3' 引物 1μL; 5' 引物 1μL; MgCl₂ 3μL; 10X 缓冲液 5μL; Taq 聚合酶 1μL (4U/μL); DDW 25μL, 首先 94℃ 解链 5min; 然后 94℃ 变性 2min, 60℃ 复性 1min, 72℃ 延伸 2min, 30 循环; 最后 72℃ 延伸 10min。RNA3 3' 端片段 X1 用引物 a 和引物 D2 (5' CCTTACTTCT-CATGGATGCTT 3', 1132 ~ 1152nt) 扩增; 5' 端片段 X2 由 E1 和引物 b (5' CCGAATTCGTAATCTTACCACTT 3') 扩增。

1.5 RNA3 的克隆和序列分析

PCR 产物经 1% 琼脂糖凝胶电泳, QIAE XII kif

(QIAGEN) 回收后, 克隆于 pGEM-T easy Vector (Promega), 转化大肠杆菌。经酶切鉴定后, 采用双脱氧法, 用 ABI PRISM 3700 DNA 自动测序仪测序。每个片段用两次独立的克隆从两端测序。利用 DNASIS 软件 (Hitachi Software Engineering Co, 1990) 对所测序列进行分析。

2 结果

2.1 CMV-Xb RNA3 的全序列测定和分析

分别对两个片段两次独立的扩增 (图 1), 酶切鉴定 (图 2) 及测序结果 (图 3) 表明, CMV-XB RNA3 全长 2205nt。总共具有两个蛋白编码阅读框架 (ORF), 其中 5' 端 (97 ~ 936nt) 编码一个 279aa 的 3a 蛋白; 3' 端 (1225 ~ 1881nt) 编码一个 218aa 的 CP 蛋白。5' 端非编码区 (NR) 为 96 nt; 基因间隔区 (IR) 长 288nt; 3' 端 NR 区域含有 324nt。

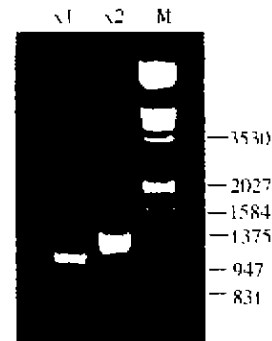


图 1 片段 X1 和 X2 的 PCR 扩增产物琼脂糖电泳
Fig.1 Agarose gel electrophoresis of PCR products of CMV-Xb Fragment X1 and X2

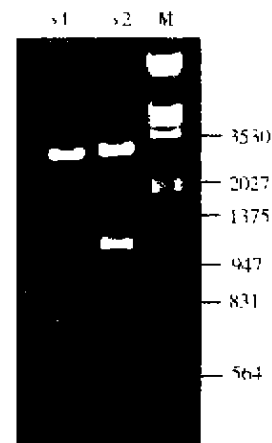


图 2 片段 X1 和 X2 重组质粒的 *EcoRI* 酶切鉴定
Fig.2 Recombinant plasmid of CMV-Xb Fragment X1 and X2 with *EcoRI* digestion

```

1   GTAATCTTACCACCTTTCTTTTCACGTCGTGCGCGTCAGTCCACGCTCTGTGTGTGTGTGTGTAGTTAGTGTGTGCTGTTTAGATTACGAAGGTTATGG
101  CTTTCCAAGGTACCAGTAGGACTTTAACTAACAGTCCCGCGCGCGTCTGACGACTTACAGAAGATATTTAGCCCGATGCCATCAAGAAGAT
201  GGCTACTGAGTGTGACCTAGGTCGACATCATGGATGCGCGGGATAACGCCATCTCTGTGACACCTCTCGTCCCAAGTAACCAGTAACAATTTATIG
301  *CTTTCTTTAAATCTGGGTATGATGCGGGTAAATGGCCTCTAAAGGCTATAATGAGCGTTCCCTCAAGTGTGTGTGCCGTTACCAGGACGGTTCCTACGG
401  ATGCTGAGGGTCTTTGAAAAATTTATTTGGCTGAOCTAGGTGATAAAGAATTATCCCAATTGATGGGCAATGTTACTTTACATAAATCATGATCTTCT
501  TGCTTTGATATCTTTCCAACCTACCTACGATTGCCCATGGAATTAATGCGCAATCGGCATCGATGTTTGGCGTAGTGTGAGAGACATGCTTATATT
601  GGTACGGTGGTACCCTGCTAGCGTGTGTAGTAAGTGGCAAGCTCAGTTTCTTCAAGAATAAATAATTACACACACCGCGTCTGGTAAGACTCTTG
701  TGTGCTTACAACAGATTAGCTGAGCATTGAAAACCGTCAGCCGTCGCTCGCCTGTTGAAGTGGCAGTTAAACAACGTTAGCTCATCGCCTAICTTIT
801  GCCGAACCTTGCTCTTAACCAAAATGCGTCTGGGACCGAGTCCGAGATTTAAACGAAAGCCCTCCCATCGCTATAGGGAGTCCGTCGCGTCCCGTAAC
901  AATAGCTTCAGATCGCAGGTGTTAACGGTCTTTAGTGTGTTTACGTTGTACTATATATACTACGTTTATCTTCGCTATGTAATAACATGAT
1001 GAGTCTAGAGTCCCGTGTGAGTTGTAACCGTAGACATCTGTGACGCGAAACCGCTGAAGATTTCCATCTGGGGTTAGTAAGTCCACATCACAGTTTFA
1101 AGGTTCAATTCCTTTGCTCCCTGTGGGCCCTTACTTCTCATGGATGCTTCCCGGAGTTAGCGTTTAGTTGTTACTTGAGTGTGCGTCTTCTT
1201 TGTGTTTGGCTCTTAGTGTATTCATGGACAAATCTGGATCTCCCAATGCTAGTAGAACCTCCGGCGTGTGCGCCCGGTAGAGGTTCTCGGTCCGCTT
1301 CTGGTGGGATGCAGGGTTGCGTGTGTTGACTCAGCAGATGCTGAGACTCAATAAAACCCTCGCCATTGGTGTGCCACTCTTAACACCCCAATCTCGT
1401 GGGTAGTGAAGCTGTAACCCCGGTTACACTTTCACATCTAATACCTGAAACCCGCTGAAATGAAAAGGGTTCATACCTTGGTAGAAGTGTGCTTTG
1501 CCAGATTGAGTACCGACTATGATAAGAAGCTGTTTCGCGCATCAAAATCAGGATTAATCCTTTCGCCAAATTTGATCTACCGTGTGGGACACAGTCT
1601 GGAAGTACCTTCATCATCGATCTTTCGTCGCGCCATCTCTGCTATGTTGGCGATGGTAACTACCAGGTTTGGTTTACAGTATGCTGCATCCCG
1701 AGTTCAGCCCAACAATAAGTTACTTTATGACCTGACCGAGATGCGCGTGTATATCGGCGACATGCGTAAGTACCGCTCTGGTCTACTGAAAGACGAT
1801 AACTAGAGAAGGACGAGATTGTAATTCATGTCGACGTCGAGCATCAACGAATTCCTATCTCAGGATGCTCCCGACTTAGTCCGTTGTTTACCGGGCT
1901 CCGAAGACGTTAAACTACACTCTCAATCCGAGTGTGAGTGGTAGTATTGCTTCAAACTGCCTGAAGTCTTAAACGTTGTTGTCACGGGAACGGG
2001 TTGTCATCCAGCTTACGGCTAAAATGGTCAGTCAATGCATCAATGCATGCGGACACTTACAAGGTTGTGAGGTACCCTTGAATCATCTCTAGATT
2101 TCTTCGGAAGGGCTTCGTGAGAAGCTGTCACGGTAATACACTGATATTACCAAGAGTGGGGTATCGCCTGTGGTTTCCACAGGTTCTCCATAAGGA
2201 GACCA
    
```

图3 CMV-Xb 的 RNA3 对应的 cDNA 序列
 编码区起始密码子用下划线标出;终止密码子用方框标出

Fig.3 The cDNA sequence of CMV-Xb RNA3

The start and end codes are showed with underline and box respectively

2.2 CMV-Xb RNA3 与其它亚组 II 株系 RNA3 序列变异分析

CMV-Xb RNA3 的全序列与其它株系(株系) RNA3 序列同源性比较如表 1 所示。

表 1 Xb 株系 RNA3 的 cDNA 序列与亚组 II 其它成员的同源性比较(%)
 Table 1 The comparison of cDNA sequence of RNA3 of Xb with those of the other strains of subgroup II (%)

株系	Xb	Q*	Trk7*	M2*	UK*	M
5'NR	-	98.9	97.9	96.7	99.0	60.8
3a 编码区	-	98.2	98.1	98.5	98.7	79.4
IR	-	95.8	95.5	95.4	96.5	69.1
CP 编码区	-	98.2	97.0	98.4	97.3	75.8
3'NR	-	93.4	94.0	94.2	94.2	64.8

* 序列来自 GenBank, Accession number, 分别为 Trk7 (L15336), M (D10539), Q (M21464), UK (Z12818), M2 (AB006813)

由表 1 可以看出, Xb 株系 RNA3 的序列和亚组 II 其它成员的序列同源性很高。各个成员之间的同源性也很高(未显示数据)。总体而言其发生变异的区域在非编码区。尤其是在基因间隔区(IR)和 3'

端 NR 区域。3a 和 CP 编码区的序列十分保守。5' NR 的变异没有较多碱基的插入和缺失现象。3a 和 CP 编码区的密码子移位情况很少。亚组 II 成员的序列长度变异只存在于 IR 区域, 有较多碱基的缺失和插入现象。与亚组 I 株系 M 的同源性不大于 80%。

2.3 与其它亚组 II 成员的 3a 和 CP 氨基酸序列比较

由图 4a 可以看出, Xb 株系编码的 3a 蛋白第 133 位出现了一个独特的 D, 而不是其它株系共有的 E。第 168 位缺失一个 I, 但在第 186 位又插入一个 I。其它位点也有氨基酸改变, 但不是其自身特有的。

由图 4b 看出, CP 的第 123 位和第 171 位分别出现了 A 和 T 两个特有氨基酸, 代替了其它株系所共有的 V 和 S。

通过 MP 氨基酸的比较可以看出, 亚组 II 成员

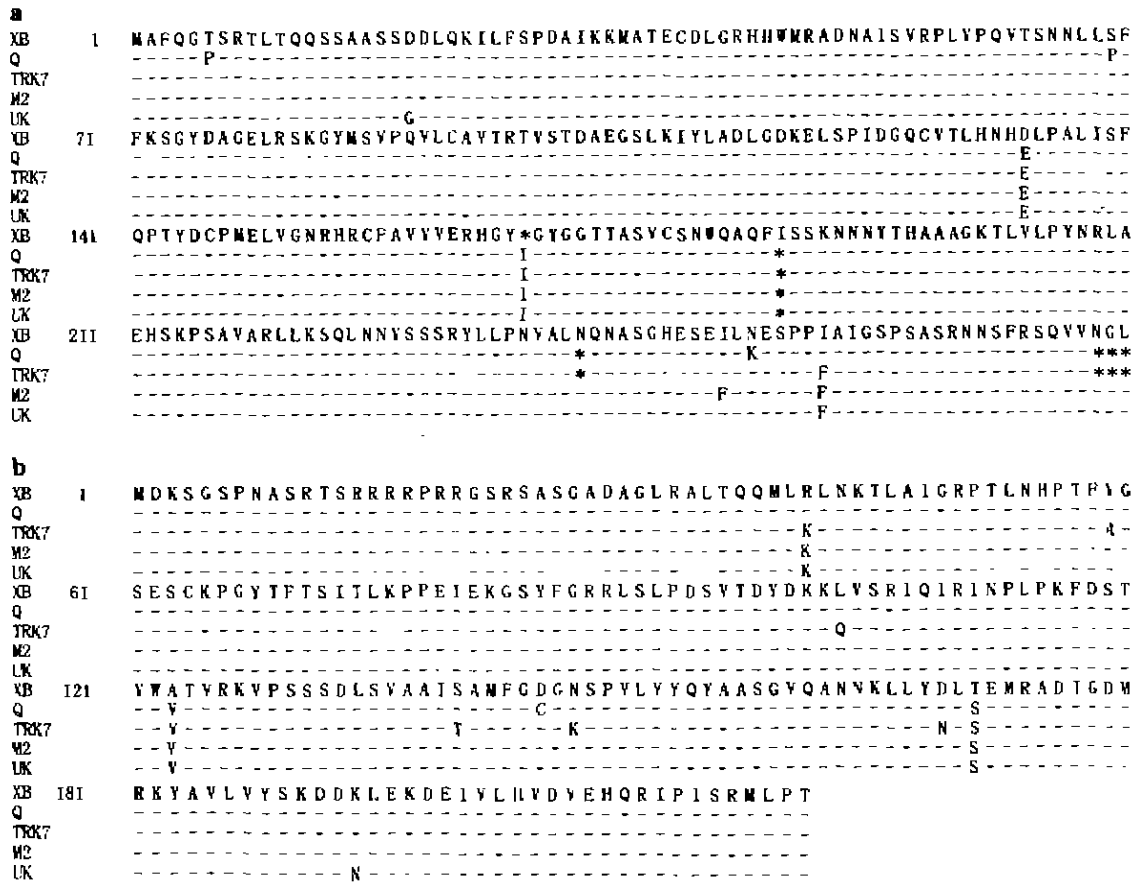


图 4 CMV-Xb 株系 3a 蛋白和 CP 氨基酸序列与亚组 II 其它成员的比较

a, 3a 蛋白氨基酸序列比较; b, CP 氨基酸序列比较; - 表示相同的氨基酸; * 表示缺失的氨基酸。

Fig. 4 The alignment of the putative amino acid sequences of 3a and CP of CMV-Xb with those of strains of subgroup II

a, The putative amino acid sequences of 3a of CMV-Xb with that of strains of subgroup II; b, The putative amino acid sequences of CP of CMV-Xb with that of strains of subgroup II; - Shows the same amino acids with the top line; * Represents the deleted amino acid compared to the top.

的 RNA3 变异也有连续性。Xb、TRK7、M2 和 UK 株系 MP 蛋白第 5 位共有的 T 变为 Q 株系中的 P; 第 259 位的 I→F。在 CP 氨基酸的序列中, 也有类似的情况。如第 41 位的 R→K。

2.4 亚组 II 成员 RNA3 的非编码区域之间序列同源性分析

表 2 亚组 II 成员间 RNA3 的非编码区域序列同源性比较

Table 2 The homology of nucleotides of the non-coding regions of the RNA3s among strains of subgroup II

株系	Xb	Q	TrK	M	UK
Xb	-	98.9	97.9	96.9	98.9
IR Q	95.8	-	97.9	95.8	100
TrK	95.5	94.0	-	95.8	97.9
Q ₁	89.6	93.3	87.5	93.7	97.9
M2	95.5	97.9	95.6	-	95.9
UK	96.5	97.5	95.3	97.8	-

株系	Xb	Q	TrK	M2	UK
Xb	-	93.1	94.0	94.2	94.2
Q		-	98.1	98.8	99.1
TrK			-	98.8	99.1
Q ₁				83.2	77.6
M2				-	99.7
UK					-

由表 2 可以看出, 在亚组 II 成员之间, 5' NR 区域的序列表现特别保守, 同源性都高于 95.5%, UK 和 Q 株系的 5' NR 区序完全相同; 在 IR 区域, 各个株系之间也高达 92.1% 以上; 在 3' NR 区, 各株系之间也高达 92.3% 以上。

3 分析与讨论

关于 CMV-Xb 株系的 CP 基因已经有过报道^[6]。但所报道的序列和我们测定的全序列中 CP 基因同源性仅为 98.6%。同一个株系的两次序列分析结

果为什么有如此大的差异?我们应用完全不同的两次克隆和两端序列分析的两次结果完全相互符合,说明没有出现株系混杂和 RNA 污染的问题。而先前的测定有可能因序列测定不准确而造成结果误差。

许多报道表明,CMV 亚组 II 株系共同在寄主上引起轻微花叶症状,在接种叶上会出现蚀纹斑。同时还可以支持卫星 RNA 的复制^[8]。我们选取不同地域的亚组 II 株系 Trk7^[9]、美国株系 UK^[10]、Q^[11]和日本株系 M2(未发表资料)作序列分析。

序列分析进一步证实,Xb 是属于亚组 II 的典型成员。亚组 II 成员 RNA3 的变异区域主要发生在 3' 非编码区域。其它区域的序列,尤其是 5' 端非编码区和 IR 区域具有和编码区相同的同源性。这是与亚组 I 成员变异特征的最明显区别。因此它们在不同寄主上或相同寄主上引起的症状比较相似,没有太明显的致病力差异。这些特征也说明亚组 II 成员在进化过程中先从 3' NR 区域发生变异,再到 5' NR、IR 和编码区域。当这些变异超过一定的阈值时,新的株系就产生了。同时可以看出,亚组 II 株系在进化过程中有重组现象发生。如 Q 株系与 UK 株系间有共同的 5' 非编码区序列,说明 Q 株系似乎与 UK 株系 RNA3 有重组关系。这些现象还需要有更多的株系序列分析后才可以进一步证实。

序列分析表明,亚组 II 成员 RNA3 序列十分保守。与亚组 I 成员相比较,无论在编码区或非编码区的变异都很小。这种极强的保守性也许就是亚组 II 成员致病力等特性彼此十分相近的根本原因。同时表明亚组 II 成员在自然界中的株系多样性没有亚组 I 成员那么丰富。

从 3a 和 CP 氨基酸序列分析发现,亚组 II 成员

之间也有变异连续性的特点。因为许多株系间有共同变异的氨基酸位点。而这些位点另外几个株系中又十分保守。

参考文献

- [1] Peden KWC, Symons RH. Cucumber mosaic virus contains a functionally divided genome[J]. *Virology*, 1973, 53: 487-492.
- [2] Palukaitis P, Roossinck P, Dietzgen HC, et al. Cucumber mosaic virus[J]. *Adv Virus Res*, 1992, 41: 281-341.
- [3] Schewinghamer MW, Symons RH. Function of cucumber mosaic virus RNA and its translation in a wheat embryo cell-free system[J]. *Virology*, 1975, 63: 252-262.
- [4] Ding SW, Anderson BJ, Huase HR et al. New overlapping gene encoded by the cucumber mosaic virus genome[J]. *Virology*, 1994, 198: 593-601.
- [5] Schewinghamer MW, Symons RH. Translation of the four minor RNA species of cucumber mosaic virus in plant and animal cell-free toad oocytes[J]. *Virology*, 1977, 79: 88-108.
- [6] 徐平东,周仲驹,林奇英,等. 黄瓜花叶病亚组 I 和 II 株系外壳蛋白基因的序列分析与比较[J]. *病毒学报*, 1999, 15: 164-171.
- [7] Mossop DW, Franki RJB, Grivell CJ. Comparative studies on tomato aspermy and cucumber mosaic viruses. V Purification and properties of a cucumber mosaic virus inducing severe chlorosis[J]. *Virology*, 1976, 74: 533-546.
- [8] Zhang L, Hanada K, Palukaitis P. Mapping local and systemic symptom determinants of cucumber mosaic virus in tobacco[J]. *J Gen Virol*, 1994, 75: 3185-3191.
- [9] Salanki K, Thole V, Balazs E, et al. Complete nucleotide sequence of the RNA3 from subgroup II of cucumber mosaic virus (CMV) strain: Trk7[J]. *Virus Res*, 1994, 31(3): 379-384.
- [10] Boccard F, Baulecombe D. Mutational analysis of cis-acting sequences and gene function in RNA3 of cucumber mosaic virus[J]. *Virology*, 1993, 193(2): 563-578.
- [11] Davies C, Symons RH. Further implications for the evolutionary relationships between tripartite plant viruses based on Cucumber mosaic virus RNA3[J]. *Virology*, 1988, 165: 216-224.