

芜菁花叶病毒杭州株系外壳蛋白基因序列分析

徐 平 许建平 薛朝阳 李德葆

(浙江农业大学生物技术研究所, 杭州 310029)

摘 要 从杭州郊区分离出一个 TuMV 强毒株系, 利用分子克隆和 Sanger 双脱氧测序方法分析了其外壳蛋白基因的 DNA 序列。该序列与已报道的其他 4 种 TuMV 分离物都具有较高的同源性, 最高达 97.6%, 最低达 89.5%, 其氨基酸序列的同源性更高, 最高达 97.9%, 最低达 94.5%。本文分析了该序列中 AT 和 GC 含量, 计算了 4 种核苷酸在有意义链中和在密码子不同位置上的出现频率, 同时还分析了该病毒外壳蛋白氨基酸遗传密码的使用频率

关键词 芜菁花叶病毒; 外壳蛋白; DNA 序列
中图分类号

Xu Ping; Xu Jianping; Xue Zhaoyang; Li Debao (Inst. of Biotech., Zhejiang Agricultural University, Hangzhou 310029, China)

Sequence analysis of TuMV coat protein gene Hangzhou strain. *Journal of Zhejiang Agricultural University*, 1994, 20(5):

Abstract One turnip mosaic virus with strong infectivity was isolated in Hangzhou suburbs. The coat protein gene was cloned and sequenced by Sanger dideoxy method. It was similar with four reported TuMV coat protein genes. The homology of nucleotides with each other was from 89.5% to 97.6% and the homology of amino acid sequences from 94.5% to 97.9%. The percentage of AT and GC, the nucleotide frequencies and the position frequencies were calculated. The amino acid usage of coat protein was analyzed.

key words turnip mosaic virus; coat protein; DNA Sequence

芜菁花叶病毒(TuMV)属马铃薯 Y 病毒组, 病毒粒子为长 720 nm, 宽 12 nm 的杆状, 其基因组长约 10kb, 编码一个约 360 kD 的多肽^[1]。它广泛分布于世界各地, 主要侵染十字花科植物, 是十字花科植物的主要病毒病病原之一。目前除以农药防治蚜虫传播外, 尚无其他有效防治手段, 从而严重地危害了农业生产, 造成巨大的经济损失。

近年来, 随着生物工程的发展, 人们利用病毒的编码基因进行植物基因工程的研究, 生产出一系列新的抗病毒植物新品种, 有效地提高了植物的抗病性。本文在克隆到 TuMV 杭州株系外壳蛋白基因的基础上, 进行了亚克隆并测定了外壳蛋白基因的全部序列, 利用计算机对该基因进行了分析。

1 材料和方法

1.1 材料

TuMV 杭州株系外壳蛋白基因由本实验室经 PCR 扩增后克隆获得, 参见文献[2]. 实验中采用的克隆载体为 pUC 18 和 pUC 19, 大肠杆菌有受体菌为 JM 109.

所有限制酶和 T4 DNA 连接酶均来自 Boehringer Mannheim 公司, 序列分析试剂盒来自 US Biochemical 公司, $[\alpha\text{-}^{35}\text{A}]\text{dATP}$ 来自 Amersham 公司, 丙烯酰胺、甲叉双丙烯酰胺和 TEMED 来自 LKB Phomacia 公司. 其他常用化学药品为国产分析纯试剂.

1.2 DNA 的提取和克隆

微量提取 DNA 采用碱裂解法^[3]. 大量提取纯化 DNA 时, 先将细菌培养过夜, 再经 $170\mu\text{g/ml}$ 氯霉素处理过夜使质粒扩增, 然后用碱裂解法裂解细菌, 并采用溴化乙锭-CsCl 梯度平衡离心法纯化获得超螺旋 DNA, 方法参见文献[3].

DNA 的亚克隆是取 $2\mu\text{g}$ 含 TuMV 外壳蛋白基因的 DNA, 经适当的限制酶切割后, 70°C 加热使限制酶失活, 加入乙醇沉淀 DNA 后再将 DNA 悬浮于水中, 与经适当限制酶切割的质粒载体混匀后, 用 T4 DNA 连接酶在 16°C 下连接过夜, 将连接产物转化 JM 109 感受态菌并涂布在 X-gal Ap 平板上. 从转化出的菌落中挑取阳性克隆, 分别微量抽提质粒进行分析, 具体方法参见文献[3].

1.3 DNA 序列分析

以纯化的质粒 DNA 为模板, 经 NaOH 变性后分别加入 pUC/M 13 上、下游通用引物, 按 Sanger 双脱氧测序法, 用 T7 DNA 聚合酶进行 DNA 序列测定反应, 反应产物经 6% 聚丙烯酰胺凝胶于 2000V 下电泳分离, 用 5% 甲醇 5% 冰醋酸固定, 将凝胶转移到滤纸上并经真空抽干, 于 Kodak X-OMAT K 光片上曝光 49 小时, 显影后读出序列, 参见文献[4], 将获得的序列输入计算机中, 利用 Genbank 和 NASA 程序分析比较所得序列和其它已报导序列的差异.

3 结果和分析

2.1 外壳蛋白基因的亚克隆

外壳蛋白基因的亚克隆按图 1 策略将含外壳蛋白基因的片段经 EcoRI 酶切后亚克隆到 pUC 18 和 pUC 19 中, 分别转化大肠杆菌并在 X-gal Ap 平板上挑取数十个白色菌落, 微量抽提质粒, 经限制酶切分析后获含 0.5kb 和 0.4kb 片段的亚克隆.

2.2 外壳蛋白基因的一级结构

利用 pUC/M 13 上下游通用引物, 以碱变性后的质粒 DNA 为模板, 通过 T7 DNA 聚合酶进行 Sanger 法双脱氧测序. 从 X 光片上读出 DNA 序列(见图 2, 命名为 TuMVCPX). 该

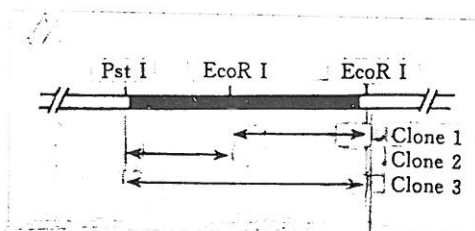


图 1 外壳蛋白基因亚克隆策略

Fig. 1. The strategy of subcloning coat protein gene

片段长度: 克隆 1 0.48kb 克隆 2 0.38kb

克隆 3 0.86kb

Fragment sizes: clone 1 0.48kb , clone 2 0.38kb ,

clone 3 0.86kb

图 2 芜菁花叶病毒外壳蛋白基因 DNA 序列

X;TuMVCPX,K;TuMVCPK,G;TuMVCPG;N;TuMVCPN;P;TuMVCP

序列与 Genbank 中存在的 4 种 TuMV(其中 2 种尚未公开发表)外壳蛋白基因的同源性都比较高,这 4 种 TuMV 外壳蛋白序列分别按作者命名为 TuMVCPK^[5]、TuMVCPG(Greenland, A. J., 未发展)、TuMVCPN^[6]和 TuMVCP(Park, Y. 等人, 未发展), 它们与 TuMVCPX 的同源性分别为 89.5%、97.6%、96.7%和 91.7%, 相对而言, 有差异的碱基大部分分布在基因的 5' 端. 其氨基酸序列见图 3, 与已报道 4 种 TuMV 外壳蛋白基因的同源性除一种为 94.5% 外, 其余均为 97.9%. 从图 2 和图 3 中可以看出由于遗传密码的兼并性, 使氨基酸序列的同源性明显大于 DNA 序列的同源性.



图 3 芜菁花叶病毒外壳蛋白氨基酸序列

Fig. 3 Amino acid sequence of TuMV coat protein

X: TuMVCPX, K: TuMVCPK, G: TuMVCPG, N: TuMVCPN, P: TuMVCP

2.3 DNA 序列的一些特性

利用计算机分析 TuMV 杭州株系外壳蛋白基因的结果表明, 双链 DNA 的 AT 出现频率为 52.13%, CG 出现频率为 47.87%. 在该基因的 DNA 序列中存在下列限制酶酶切位点: Alu I (563, 645, 650), Bcl I (478), Dde I (646), EcoR I (488), Hae III (359, 545), Hinf I (208), Mbo I (265, 298, 478, 495), Mbo I (70, 152, 611), MluI (657), Taq I (264, 652) (数字为酶切位点位置). 在有意义单链 DNA 中各种核苷酸的出现频率为 A 33.10%, C 20.99%, G 26.87%, T 19.03%, A 的出现频率明显高于 T. 各种核苷酸在密码不同位置的出现频率见表 1, 由表 1 可知 A

表 1 阅读框中 4 种核苷酸在不同位置的出現頻率

Table 1 Position frequency of 4 nucleotides in open reading frame

	A	C	G	T
1	96	58	98	87
2	112	57	50	70
3	79	67	85	58

在密码中第二个位置的出现频率较高, T 在第一个位置的出现频率较低. 表 2 为外壳蛋白基因遗传密码的使用频率, 与植物和动物遗传密码的使用频率比较, TuMV 外壳蛋白基因与双子叶植物较为相近, 而与人差异较大.

表 2. 外壳蛋白的氨基酸遗传密码使用频率

Table 2 Codon usage in coat protein genes of turnip mosaic virus

AA	Codon	TuWV		Plant ^① (%)		Human ^①
		No.	AA (%)	Dicots	Monocots	(%)
G	GGG	2	11.1	12	21	24
	GGA	5	27.8	38	17	17
	GGT	6	33.3	34	18	15
	GGC	5	27.8	16	44	44
E	GAG	14	53.8	51	75	60
	GAA	12	46.2	49	25	40
D	GAT	8	42.1	58	27	38
	GAC	11	57.9	42	73	62
V	GTG	4	28.6	29	36	50
	GTA	5	35.7	12	8	9
	GTT	4	7.1	39	19	13
	GTC	1	28.6	20	37	27
A	GCG	3	14.3	6	22	12
	GCA	11	52.4	25	16	17
	GCT	4	19.0	42	24	31
	GCC	3	14.3	27	38	40
R	AGG	3	13.6	25	26	23
	AGA	6	27.3	30	9	24
	CGG	1	4.5	5	13	15
	CGA	5	22.7	8	4	10
	CGT	4	18.3	21	12	9
	CGC	3	13.6	11	36	19
S	AGT	3	30.0	14	8	11
	AGC	2	20.0	18	26	29
	TCG		0.0	6	14	7
	TCA	1	10.0	19	11	11
	TCT	2	20.0	25	15	17
	TCC	2	20.0	18	26	26
K	AAG	13	65.0	61	86	55
	AAA	7	35.5	39	14	45
N	AAT	5	27.8	45	25	34
	AAC	13	72.2	55	75	66
M	ATC	13	100	100	100	100
	ATA	3	30.0	18	11	13
	ATT	4	40.0	45	24	23
	ATC	3	30.0	37	65	64
T	ACG	8	38.1	8	21	12
	ACA	6	28.6	27	14	21
	ACT	3	14.3	35	19	20
	ACC	4	19.0	30	46	47
W	TGG	3	100	100	100	100
END	TGA	1	100	33	34	—
	TAG	0	—	19	36	—
	TAA	0	0.0	48	30	—
C	TGT		0.0	44	30	30
	TGC	1	100	56	70	70
Y	TAT	2	25.0	43	21	47
	TAC	6	75.0	57	79	53
L	TTG	4	16.7	26	14	0
	TTA	6	24.9	10	3	5
	CTG	3	12.5	9	28	46
	CTA	3	12.5	8	9	7
	CTT	4	16.7	28	15	11
	CTC	4	16.7	19	31	22
F	TTT	4	44.4	45	25	35
	TTC	5	55.6	55	75	65
Q	CAG	10	66.7	41	46	26
	CAA	5	33.3	59	54	74
H	CAT	4	66.7	54	33	42
	CAC	2	33.3	46	67	58
P	CCG	4	40.0	9	23	11
	CCA	3	30.0	42	34	24
	CCT	1	10.0	32	17	24
	CCC	2	20.0	17	26	41

注: ①见参考文献[7]

3 讨 论

经分析表明 TuMV 杭州株系外壳蛋白基因与其他 4 种已经报道的外壳蛋白基因都具有较高的同源性, 从而确证了该基因可编码 TuMV 的外壳蛋白, 为下一步转基因植物工作打下基础. 但同时在核酸水平上杭州株系又与其他分离物具有一定的区别, 在 DNA 中这些区别大

多分布在 5' 端, 在氨基酸序列中则大多出现在 N 端。由于氨基酸密码子的兼并性, 使氨基酸序列之间的同源性比 DNA 之间的更高。比较 DNA 和氨基酸序列表明, 在分子水平上, 不同的 TuMV 株系之间可能具有一定的差异, 因此利用这种差异, 通过分子生物学手段可迅速准确地划分出不同的株系, 从而减少人们利用生物学鉴定划分株系所花费的大量人力和时间。目前人们已通过生物学鉴定方法和单克隆抗体分离划分出许多种 TuMV 株系^[8,9], 这些不同的 TuMV 株系在致病力上存在着较大的差异, 而这种差异都会反映在 DNA 水平上, 通过比较不同株系之间致病力和 DNA 序列的差别, 就可能推断出与致病有关的基因, 从而认识抗病机制提供理论依据。

利用分子生物学手段进行基因工程具广阔的应用前景, 它为人们改造植物、产生新的优良品种提供了新的途径。但我们也应意识到, 由于目前许多分子生物学方面的基本理论问题尚未彻底解决, 产生新的优良品种并非一蹴而就, 仍需要进行大量艰苦的工作。

参考文献

- 1 Yurchot J, Koonin EV, Carrinton JC. The 35 kDa protein from the N-terminus of the polyprotein functions as a third virus-encoded proteinase. *virology*, 1991, 185: 527-535.
- 2 徐平, 许建平, 陈集双, 张耀洲, 李德葆. 芜菁花叶病毒外壳蛋白的序列分析及在大肠杆菌中表达. *病毒学报*, 1994. 印刷中.
- 3 李德葆, 徐平主编. 重组 DNA 的原理和方法. 浙江: 浙江科技出版社, 1994.
- 4 Ausubel FM, *et al.* Short protocols in molecular biology. 2nd ed. New York: John Wiley & Sons, 1992.
- 5 孔令洁. 芜菁花叶病毒 CP 基因的克隆及其植物表达载体的构建. *中国科学 B 辑*, 1992, 4: 374-380.
- 6 Nicolas O, Laliberte JF. The complete nucleotide sequence of turnip mosaic potyvirus RNA. *J. Gen. Virol.*, 1992, 73: 2785-2793.
- 7 洪益国, 王小凤, 田波. 印度木薯花叶病毒外壳蛋白的研究. I. 计算机分析外壳蛋白的一些特性. *微生物学报*, 1993, 33(4): 255-259.
- 8 刘翔平. 我国十省(市)十字花科蔬菜芜菁花叶病毒(TuMV)株系分化研究. *病毒学杂志*, 1990, 1: 82-87.
- 9 刘翔平. 芜菁花叶病毒单克隆抗体杂交瘤细胞株的建立及对 7 个株系原决定簇的分析. *科学通报*, 1991, 18: 1414-1417.