キュウリモザイクウイルスY系統RNA3の全塩基配列決定ならびにQ系統との比較

誌名	日本植物病理學會報 = Annals of the Phytopathological Society of Japan
ISSN	00319473
著者	新田, 直人 増田, 税 桑田, 茂 高浪, 洋一
巻/号	54巻4号
掲載ページ	p. 516-522
発行年月	1988年10月

日植病報 54: 516–522 (1988)
Ann. Phytopath. Soc. Japan 54: 516–522 (1988)

Comparative Studies on the Nucleotide Sequence of Cucumber Mosaic Virus RNA3 between Y Strain and Q Strain

Naoto NITTA*, Chikara MASUTA**, Shigeru KUWATA**
and Yoichi TAKANAMI**

Abstract

The complete nucleotide sequence of cucumber mosaic virus RNA3 of Y strain has been determined and was compared with the reported sequence of RNA3 of Q strain. The sequence of Y strain provided the amino acid sequences of two proteins: the 5′-terminal 3A protein with 279 amino acids and the 3′-terminal viral coat protein with 218 amino acids, while 3A and coat proteins of CMV-Q have been reported to have 333 and 236 amino acids, respectively. The 3′-terminal region of Y strain can be arranged into the secondary structures similar to those proposed for that of Q strain, although there are considerable differences in sequence.

(Received May 16, 1988)

Key words: cucumber mosaic virus, nucleotide sequence, RNA3.

INTRODUCTION

Cucumber mosaic virus (CMV) has a single-stranded (+) RNA genome composed of RNA1, 2, and 3[7,10]. RNA3 encodes 3A and coat proteins, and subgenomic RNA4 is a mRNA for coat protein, generated from RNA3 during replication[1,13]. The complete nucleotide sequences of RNA1[12], RNA2[11], and RNA3[1] of CMV-Q strain and a sequence near the 5′ terminus of RNA4 of CMV-Y strain[5] have been reported. Comparison of nucleotide sequences among different strains of a virus provides us various useful information on the essential structure(s) of the molecule related to biological functions of the viral RNAs, such as regulation of replication, translation or generating of subgenomic RNA and on phylogenetic relation. Thus, we have determined the complete nucleotide sequence of RNA3 of Japanese CMV-Y strain[17]. A part of this work was presented at the Annual Meeting of the Phytopathological Society of Japan in 1988.

MATERIALS AND METHODS

RNA purification. CMV-Y was propagated in tobacco plants (*Nicotiana tabacum* L. cv. Xanthi nc) and viral RNA was extracted as described[15]. RNA3 was purified by two cycles of sucrose density-gradient centrifugation[16].

cDNA cloning. The RNA was polyadenylated and its double-stranded cDNA was synthesized according to Gubler and Hoffman[2], and then ligated into *Sma*I site of plasmid pUC119 (Takara Shuzou Co., Ltd., Japan). Two clones of pY3-46 and pY3-79, each of which

* Yokohama Center, Applied Plant Research Laboratory, Japan Tobacco Inc., 6–2, Umegaoka, Midori-ku, Yokohama 227, Japan 日本たばこ産業㈱植物開発研究所横浜センター
** Life Science Research Laboratory, Japan Tobacco Inc., 6–2, Umegaoka, Midori-ku, Yokohama 227, Japan 日本たばこ産業㈱生命科学研究所

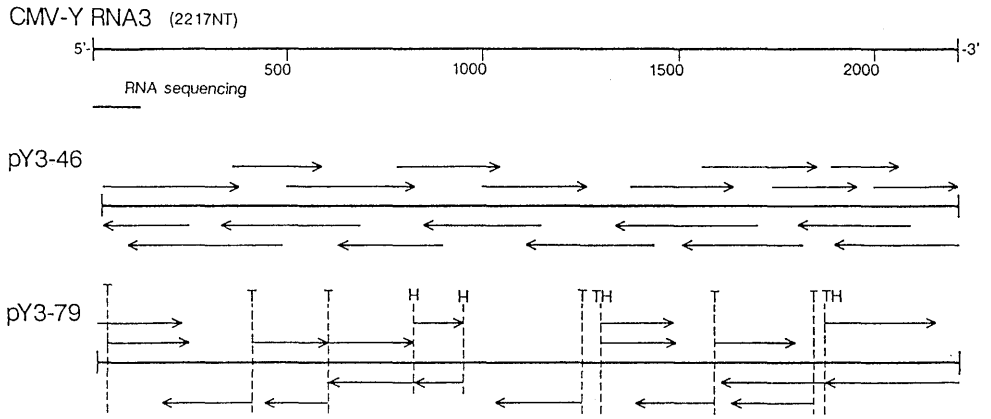


Fig. 1. Sequencing strategy. T and H indicate *Taq*I and *Hin*cII sites of pY3-79 insert, respectively. Arrows indicate the direction and extent of sequencing.

contained an about 2.2 Kbp insert and hybridized with RNA3 by Southern hybridization, were selected. pY3-46 insert was subcloned into M13 phage, and deleted using Exonuclease III[4,18]. pY3-79 insert and its fragments digested by *Taq*I or *Hin*cII were subcloned into M13 phage.

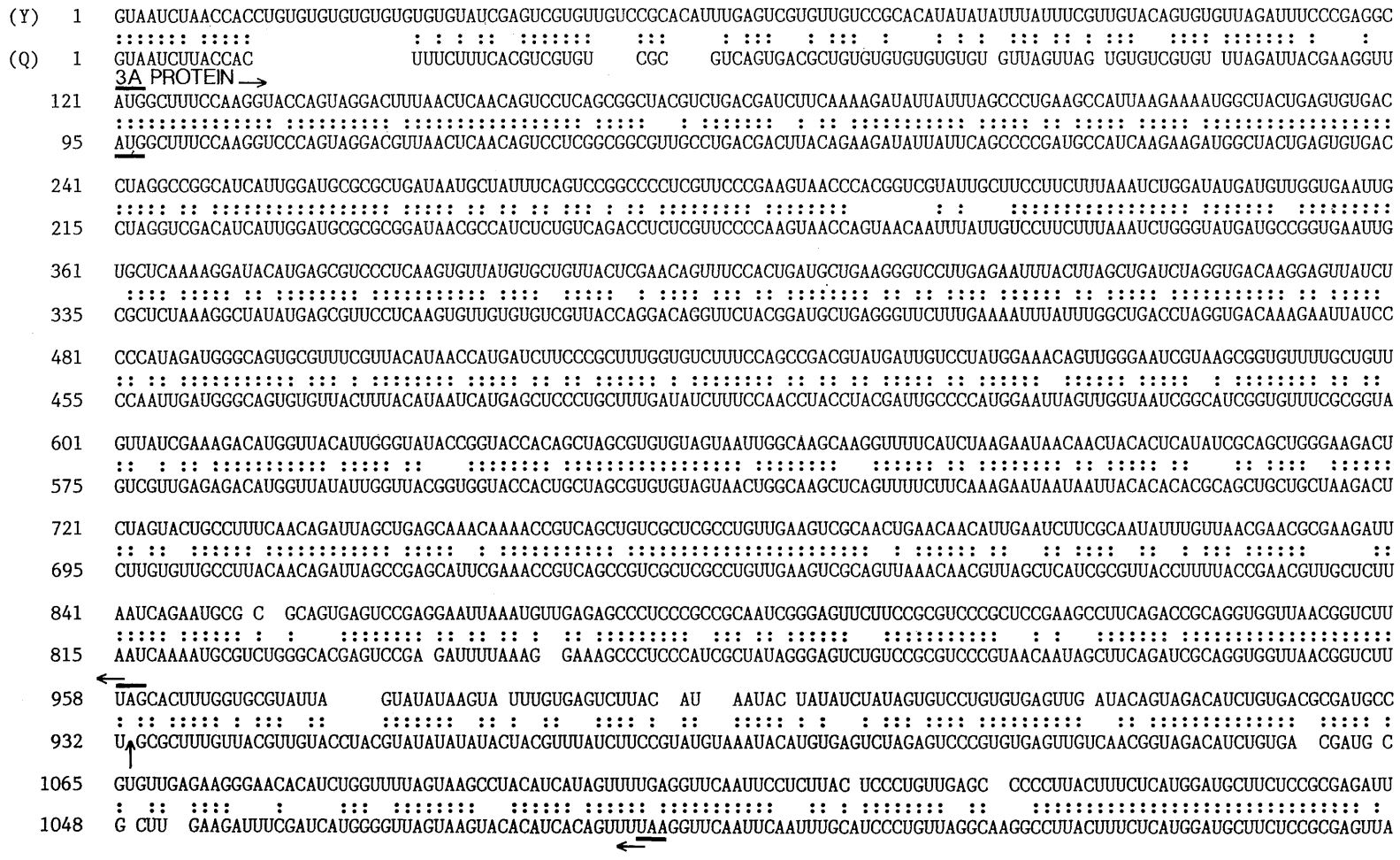
Sequencing. Sequencing by dideoxy chain termination method[9] was carried out on both strands as shown in Fig. 1. The reactions were conducted at 50 C[8]. The sequence at the 5′ terminus of RNA3 was determined by dideoxy method using a synthetic oligodeoxyribonucleotide (5′-ACTGGTACCTTGGAAAG-3′), complementary to nucleotides 125–141 to prime reverse transcription in the presence of dideoxynucleotides[8].

RESULTS AND DISCUSSION

CMV-Y RNA3 turned out to be 2217 nucleotides long although that of Q strain has been reported to have 2193 nucleotides[1], and over all sequence homology between RNA3 of Y strain and that of Q strain was 73% (Fig. 2). Our data provided the amino acid sequences of the two proteins coded for by the RNA: the 5′-terminal 3A protein with 279 amino acids (Mr 30,478) and the 3′-terminal viral coat protein with 218 amino acids (Mr 24,113). Predicted amino acid composition of the coat protein is in good agreement with its amino acid composition determined empirically[14] (Table 1). On the other hand, 3A and coat proteins of CMV-Q have been reported to have 333 and 236 amino acids, respectively[1] (Figs. 3A and 4A).

If we assume an insertion of A between nucleotides 932–933 in the reported sequence of Q strain, the open reading frame of 3A protein ends with an UAG termination codon at the inserted position and the number of its amino acid residues would be 279. In addition, the C-terminal amino acid sequence becomes homologous to that of Y strain (Fig. 3B) and homology of 3A protein between both strains is raised up from 67% to 82%. Because 3A protein itself has not been obtained as a material which enables us to analyze its physical and chemical properties, at present it is difficult to discuss the discrepancy in the length of the open reading frames of both strains.

If we assume one nucleotide insertion between nucleotides 1815–1816 in Q strain, a UAG termination codon of coat protein appears at positions 1870–1872; the number of its amino acid residues would be 217. The C-terminal region becomes homologous to that of Y strain (Fig. 4B) and homology of amino acid sequence between both strains increases from 67% to 78%. Moreover, the amino acid composition based on the insertion does not contradict that obtained by direct amino acid analysis[1] (Table 1) and its molecular weight (Mr) becomes 24,004 which


(Y) 1
(Q) 1
3A PROTEIN
121
95
241
215
361
335
481
455
601
575
721
695
841
815
958
932
1065
1048

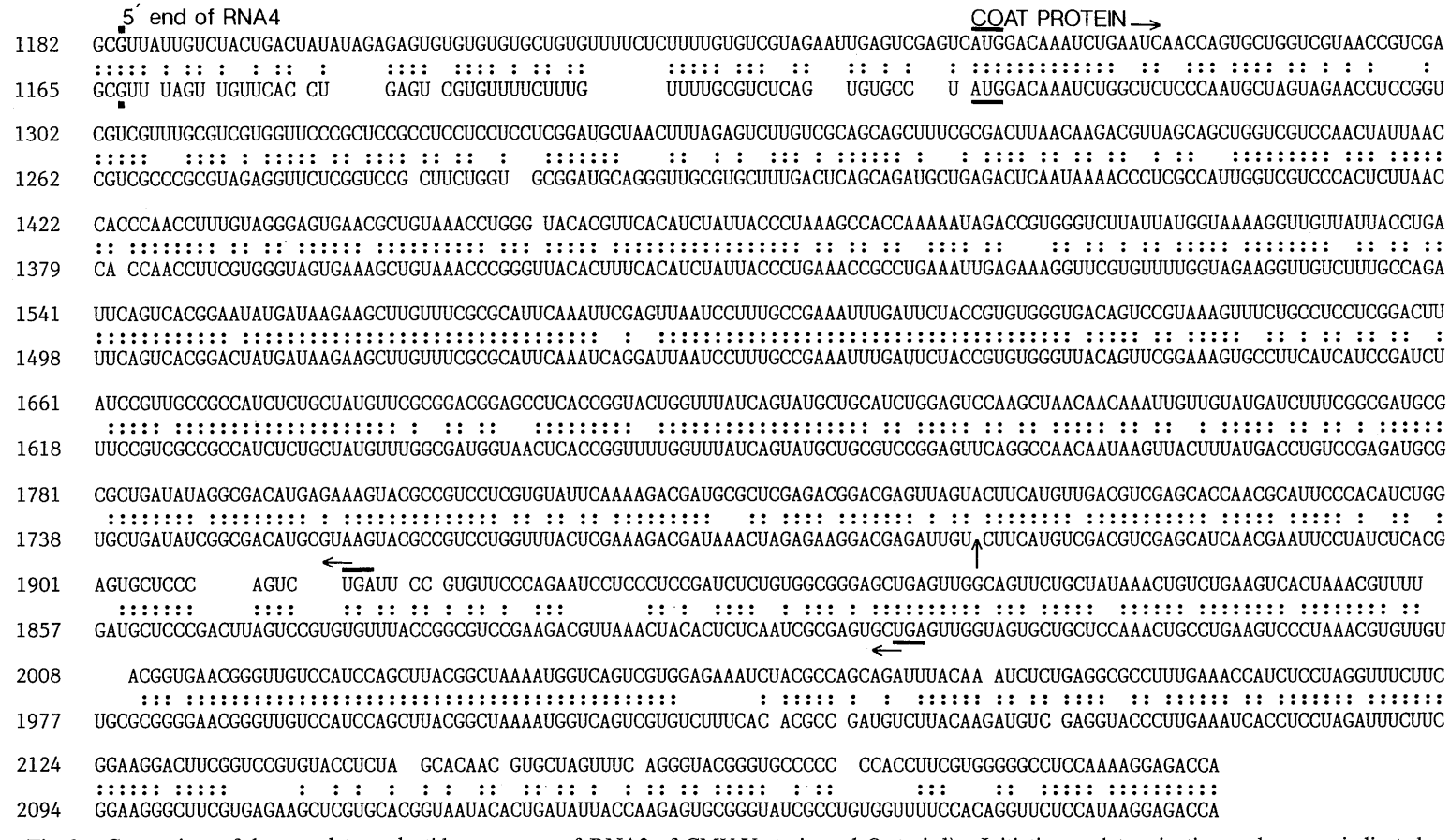


Fig. 2. Comparison of the complete nucleotide sequences of RNA3 of CMV-Y strain and Q strain[1]. Initiation and termination codons are indicated by (—). The positions of nucleotides 932–933 and 1815–1816 of Q strain are indicated by (↑).

Table 1. Amino acid composition of CMV-Y and -Q coat proteins

Amino acid (A.A.)	Y strain		Q strain		
	from sequence	from A.A. analysis[14]	from sequence[1]	from sequence with insertion	from A.A. analysis[1]
Ala	18 (8.3)[a]	(8.5)	15 (6.4)	15 (6.9)	(6.4)
Arg	20 (9.2)	(8.0)	20 (8.5)	18 (8.3)	(8.5)
Asn+Asp	22 (10.1)	(10.3)	25 (10.6)	23 (10.6)	(10.2)
Cys	1 (0.5)	(0.0)	4 (1.7)	1 (0.5)	(0.4)
Gln+Glu	12 (5.6)	(5.7)	13 (5.5)	13 (6.0)	(6.8)
Gly	11 (5.0)	(4.9)	14 (5.9)	13 (6.0)	(5.9)
His	3 (1.4)	(1.4)	2 (0.8)	3 (1.4)	(1.7)
Ile	8 (3.7)	(3.7)	10 (4.2)	11 (5.1)	(4.7)
Leu	21 (9.6)	(9.9)	22 (9.3)	20 (9.2)	(9.3)
Lys	13 (6.0)	(6.2)	15 (6.4)	14 (6.5)	(6.4)
Met	4 (1.8)	(1.7)	6 (2.5)	6 (2.8)	(2.1)
Phe	5 (2.3)	(2.3)	6 (2.5)	4 (1.8)	(2.5)
Pro	11 (5.0)	(5.7)	14 (5.9)	14 (6.5)	(6.8)
Ser	27 (12.4)	(11.5)	29 (12.3)	24 (11.1)	(10.6)
Thr	12 (5.5)	(5.3)	13 (5.5)	11 (5.1)	(5.9)
Trp	1 (0.5)	(0.5)	2 (0.8)	2 (0.9)	(0.4)
Tyr	9 (4.1)	(4.4)	9 (3.8)	7 (3.2)	(3.8)
Val	20 (9.2)	(10.0)	17 (7.2)	18 (8.3)	(7.6)
Total	218		236	217	
Mr	24,113		26,200	24,004	

a) mol/mol protein (mol/100 mol)

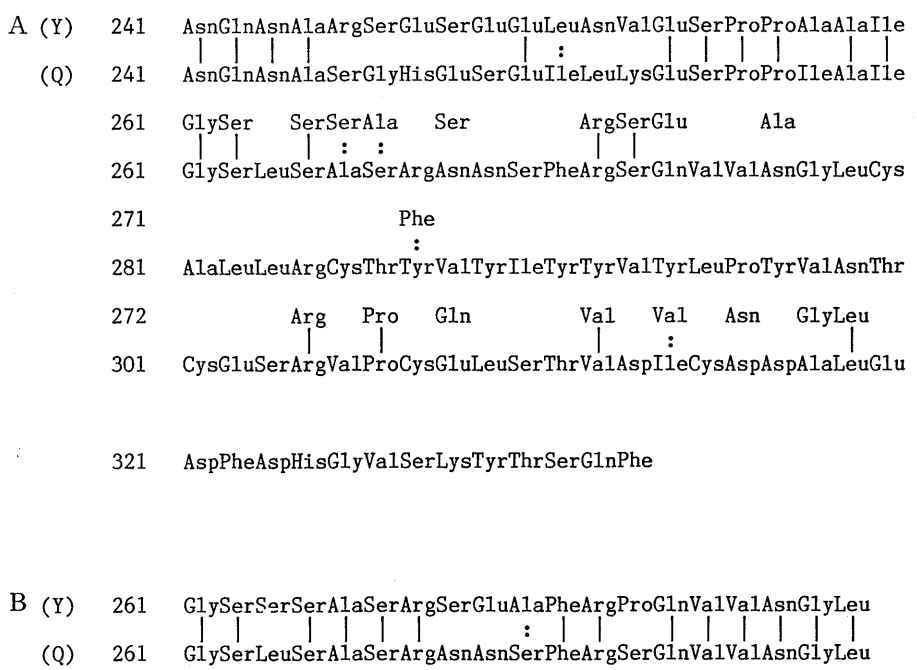


Fig. 3. Comparison of the C-terminal amino acid sequence of 3A protein of Y strain with that of original Q strain[1] (A) or with that of Q strain modified based on an insertion of A between nucleotides 932–933 (B). Identical residues (|) and similar residues (:) are according to Dayhoff classification.

```
A (Y)  180  MetArgLysTyrAlaValLeuValTyrSerLysAspAspAlaLeuGluThrAspGluLeu
  (Q)  179  MetArgLysTyrAlaValLeuValTyrSerLysAspAspLysLeuGluLysAspGluIle

       200  Val   LeuHis         Val   Asp   ValGluHisGln     ArgIle
       199  ValPheMetSerThrSerSerIleAsnGluPheLeuSerHisGlyCysSerArgLeuSer

       211  ProThrSer         Gly      Val         Leu     ProVal
       219  ProCysValTyrArgArgProLysThrLeuAsnTyrThrLeuAsnArgGluCys

B (Y)  200  ValLeuHisValAspValGluHisGlnArgIleProThrSerGlyValLeuProVal
  (Q)  199  ValLeuHisValAspValGluHisGlnArgIleProIleSerArgMetLeuProThr
```

Fig. 4. Comparison of the C-terminal amino acid sequence of coat protein of Y strain with that of original Q strain[1] (A) or with that of Q strain modified based on one nucleotide insertion between nucleotides 1815–1816 (B).

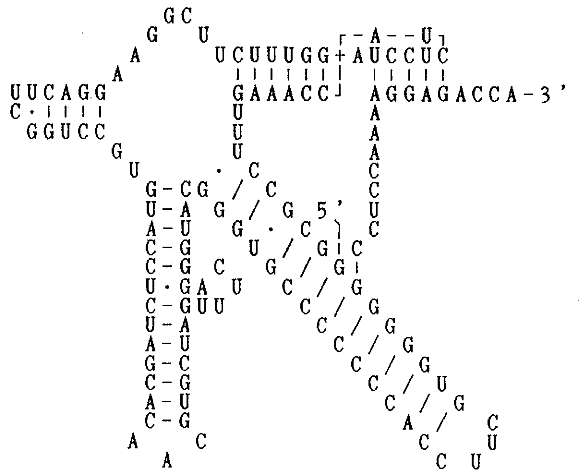


Fig. 5. Secondary structure of 3′-terminal region of CMV-Y RNA3 modeled after that of Q strain proposed by Joshi *et al.*[6]

is rather closer to Mr 24,500 estimated by SDS-polyacrylamide gel electrophoresis[3] than Mr 26,200 predicted from the original sequence. A converse assuming of a deletion of A at position 1859 of Y strain, which corresponds to the insertion site of Q strain, leads to an unlikely long coat protein (Mr 30,878), and furthermore, putative C-terminal region is not homologous to that of Q strain. The nucleotide sequence around residues 1815–1816 of Q strain was determined by sequencing only one clone on one strand[1], while the cloning and sequencing procedures used here for the corresponding position of Y strain ensured maximum confidence in the deduced sequence. The origin of the discrepancy found in the C-terminal regions of viral coat protein might be arisen from a mistake in sequence of CMV-Q or an evolutionary divergence. However, it seems unlikely that either of the same viruses, even though different in strains, has such a long deletion or insertion of 18 amino acid residues in their coat proteins, which should induce substantial morphological changes of virus particles.

Wilson *et al.*[19] and Joshi *et al.*[6], independently, proposed secondary structure models of 3′ termini of CMV-Q RNAs. In spite of considerable differences in sequence, 3′-terminal region of CMV-Y RNA3 can also be arranged into similar models of the secondary structure (Fig. 5).

Literature cited

1. Gould, A.R. and Symons, R.H. (1982). Eur. J. Biochem. 126: 217–226.
2. Gubler, U. and Hoffman, B.J. (1983). Gene 25: 263–269.
3. Habili, N. and Francki, R.I.B. (1974). Virology 57: 392–401.
4. Henikoff, S. (1984). Gene 28: 351–359.
5. Hidaka, S., Tsunasawa, S., Yoon, J., Narita, K., Takanami, Y., Kubo, S. and Miura, K. (1985). J. Biochem. 97: 161–171.
6. Joshi, R.L., Joshi, S., Chapeville, F. and Haenni, A.L. (1983). EMBO J. 2: 1123–1127.
7. Lot, H., Marchoux, G., Marrou, J., Kaper, J.M., West, C.K., Van Vloten-Doting, L. and Hull, R. (1974). J. gen. Virol. 22: 81–93.
8. Masuta, C., Zuidema, D., Hunter, B.G., Heaton, L.A., Sopher, D.S. and Jackson, A.O. (1987). Virology 159: 329–338.
9. Mizusawa, S., Nishimura, S. and Seela, F. (1986). Nucleic Acids Res. 14: 1319–1324.
10. Peden, K.W.C. and Symons, R.H. (1973). Virology 53: 487–492.
11. Rezaian, M.A., Williams, R.H.V., Gordon, K.H.J., Gould, A.R. and Symons, R.H. (1984). Eur. J. Biochem. 143: 277–284.
12. Rezaian, M.A., Williams, R.H.V. and Symons, R.H. (1985). *Ibid.* 150: 331–339.
13. Schwinghamer, M.W. and Symons, R.H. (1977). Virology 79: 88–108.
14. Takanami, Y. (1977). Thesis of Kyushu Univ. Japan.
15. Takanami, Y. (1981). Virology 109: 120–126.
16. Takanami, Y. and Fraenkel-Conrat, H. (1982). *Ibid.* 116: 372–374.
17. Tomaru, K. and Hidaka, Z., (1960). Bull. Hatano Tobacco Exp. Sta. 46: 143–149.
18. Yanisch-Perron, C., Vieira, J. and Messing, J. (1985). Gene 33: 103–119.
19. Wilson, P.A. and Symons, R.H. (1981). Virology 112: 342–345.

和 文 摘 要

新田直人・増田 税・桑田 茂・高浪洋一：キュウリモザイクウイルス Y 系統 RNA3 の全塩基配列決定ならびに Q 系統との比較

キュウリモザイクウイルス Y 系統 RNA3 の全塩基配列を決定し，既報の Q 系統 RNA3 の塩基配列と比較した。その結果，Y 系統の塩基配列から 5′ 側にコードされている 3A 蛋白は 273 アミノ酸，また 3′ 側にコードされているウイルス外被蛋白は 218 アミノ酸から成ると予想され，Q 系統における 3A および外被蛋白のアミノ酸残基数 333 および 236 とは大幅に相違した。両系統の 3′ 末端領域は，塩基配列にかなりの相違があるにもかかわらず，類似の 2 次構造を取りうることが判明した。