한국식물병리학회지 11(3) : 271-277(1995)
Korean J. Plant Pathol. 11(3) : 271-277(1995)

Nucleotide Sequence Analysis and Secondary Structure Modeling of the 3'-Noncoding Regions of Two Korean Strains of Turnip Mosaic Virus

Jang Kyung Choi*, Ki Hyun Ryu[1], Gug Seoun Choi and Won Mok Park[1]
Department of Agricultural Biology, College of Agriculture, Kangwon National University, Chuncheon 200-701, Korea
[1]Department of Agricultural Biology, College of Natural Resources, Korea University, Seoul 136-701, Korea

순무 모자이크 바이러스 두 한국계통의 3' 말단 비번역부위에 대한 염기서열분석 및 2차구조 모델링

최장경*. 류기현[1]. 최국선. 박원목[1]
강원대학교 농과대학 농생물학과, [1]고려대학교 자연자원대학 농생물학과

ABSTRACT : The RNA nucleotide sequences of the 3'-noncoding regions (3'-NCRs) of two Korean strains of turnip mosaic virus (TuMV), Ca and cqs, have been determined from their cDNA clones that encompassed the 3'-terminal regions of the viral genomic RNAs. The 3'-NCRs of both strains were 209 nucleotides long, terminated with GAC residues and poly (A) tails. The potential polyadenylational signal motif, UAUGU, was located 140 nucleotides upstream from the poly (A) tail in each of the virus. A highly conserved hexanucleotide sequence [A G U G A/U G/C], which was common in the 3'-NCRs of the potyvirus RNAs, was also found at the regions of 119 bases upstream from the 3'-end. Comparison of the 3'-NCRs of the two Korean isolates with those of four strains from Canada, China and Japan showed significantly identical genotypes (94.3~99.5%). The secondary structure of three loops with long stems was found within the 3'-NCRs by sequence analysis. The substituted bases in the region among the six TuMV strains did not alter their secondary structures. Length of the 3'-NCRs of the known 11 potyviral RNAs and TuMV RNAs was different from one another and their nucleotide sequences showed 55.7% to 24.0% of homology. The 3'-NCR, therefore, is considered to be useful for phylogenetic studies in potyviruses.

Key words **: turnip mosaic virus, two Korean isolates (Ca, cqs), potyvirus, nucleotide sequence analysis, 3'-noncoding region (3'-NCR), secondary structure modeling.**

Turnip mosaic virus (TuMV) is a definite species of the genus *Potyvirus* in the taxonomic family Potyviridae of plant viruses. TuMV causes diseases on economically important vegetable crops in Korea, particularly Chinese cabbage (*Brassica campestris* L. ssp. *pekinensis*) and radish (*Raphanus sativus*).

The genome organization of potyvirus has been well characterized (3, 12, 21). The genome of potyvirus is composed of a positive-sense ssRNA of about 10 kb, which is linked covalently at its 5'-end to a virus-encoded protein (VPg) and polyadenylated at its 3'-end. The genomic RNA is translated into a large polyprotein precursor that is proteolytically processed into at least eight translation products by three proteinases encoded by virus itself (3).

Partial nucleotide sequences of genomic RNAs in several strains of TuMV have been reported (5, 11, 15, 24). Recently, the complete nucleotide sequence and genome organization of TuMV has been reported (17).

Although many serological assays are useful for

*Corresponding author.

virus detection, several potyviruses share serological similarities, so that they may not be distinguishable from one another (13). When unique regions such as 3'-noncoding region (3'-NCR) are targeted, specific viruses and even strains of a virus can be distinguished (7, 8, 25).

We have isolated genome RNAs and cloned the 3'-terminals including the partial nuclear inclusion body (NIb) and entire coat protein (CP) gene from two Korean TuMV strains, as a step toward developing molecular probes for virus detection and genetically engineered virus resistant plants (5, 6, 18, 22, 23). Here, we report the nucleotide sequences of the 3'-NCRs of genome RNAs of two Korean TuMV strains and compared them with those of other four TuMV isolates (11, 15, 17) and other 11 potyviruses (1, 3, 4, 8, 9, 10, 12, 16, 19, 20, 21). We also predict secondary structure within the 3'-NCR of the viral RNA and discuss the structural conservation within different strains.

MATERIALS AND METHODS

Virus sources and their nucleotide sequences. Two strains of TuMV, designated as TuMV-Ca and TuMV-cqs, were originally obtained from naturally infected Chinese cabbage leaves showing severe mosaic and small black necrotic spots in the high altitude area of Daekwallyeong, Kangwon-Do, in Korea (6, 18). Synthesis and cloning of cDNAs of the two isolates were already reported previously (5, 22). The 3'-end of TuMV-Ca was sequenced using pTUCA35 and confirmed using pTUCA31 which has overlapping region of pTUCA35 (22). pTUS6 was used as a parent clone for determining nucleotide sequence of the TuMV-cqs (5). Nucleotide sequences of the 3'-NCRs of the two Korean TuMV strains (TuMV-Ca, EMBL X79366; TuMV-cqs, EMBL X83968) and four foreign TuMV strains from Canada (TuMV-CAN, EMBL D10927) (17), China (TuMV-CH, EMBL X52804) (11) and Japan (TuMV-JA-1 and TuMV-JA-31) (15) were used for primary and secondary structural analyses.

Nucleotide sequences of 3'-NCR from the following viruses of the genus *Potyvirus* were used for comparison : *Kalanchoë* mosaic virus (KMV) (10), papaya ringspot virus (PRV) (19), peanut stripe virus (PStV) (4), plum pox virus (PPV) (12), potato virus Y-N (PVY-N) (21), soybean mosaic virus-N (SMV-N) (8), sweet potato feathery mottle virus (SPFMV) (1), tobacco etch virus (TEV) (3), watermelon mosaic virus 2 (WMV 2) (8), wheat streak mosaic virus (WSMV) (16) and zucchini yellow mosaic virus (ZYMV) (9).

Sequence analysis. The nucleotide sequences of the 3'-NCRs of TuMV-RNAs of two Korean strains were compared with other 4 TuMV strains from foreign countries. Sequence data were compiled and analyzed by the multiple sequence alignments, phylogenetic relationships and Zuker's RNA secondary structure programs of the PC/GENE Software Version 6.6 (IntelliGenetics, Inc.). Modeling of secondary structure was analyzed with the program described by Abrahams *et al.* (2).

RESULTS

Determination of nucleotide sequence of the 3'-NCR. The RNA nucleotide sequences of the 3'-NCRs of two Korean strains, Ca and cqs, have been determined from their cDNA clones that encompassed the 3'-terminal region of the viral genomic RNAs. The 3'-NCRs of the two Korean strains were 209 nucleotides long, terminated with GAC residues and poly (A) tails of 19 residues for TuMV-Ca and 15 residues for TuMV-cqs. Multiple alignments of the nucleotide sequences between two Korean strains and four foreign strains are shown in Fig. 1. This 3'-NCR region was the most conserved among the TuMV strains with sig-

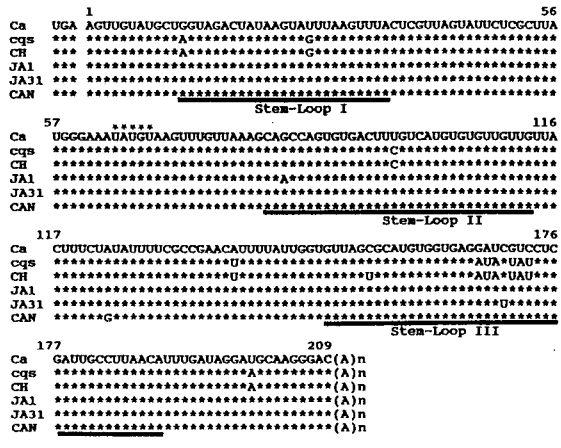

Fig. 1. Multiple alignments of the 3'-noncoding regions of the nucleotide sequences between two Korean strains (Ca and cqs) and four foreign strains of turnip mosaic virus (TuMV). Nucleotides identical to those of TuMV-Ca are shown by asterisk. Nucleotide sequences involved in formation of stem-loop structures are underlined.

Table 1. Percentages of nucleotide sequence similarities in the coat protein (below diagonal) and the 3'-noncoding regions (above diagonal) among six geographically distinct turnip mosaic virus (TuMV) strains[a]

TuMV strains	% nucleotide sequence similarity					
	Ca	cqs	CH	CAN	JA-1	JA-31
Ca	–	94.8	94.4	99.5	99.5	99.5
cqs	94.1	–	99.5	94.4	94.4	95.8
CH	94.4	97.6	–	93.9	93.9	95.3
CAN	95.1	94.8	94.4	–	99.1	99.1
JA-1	96.9	96.5	96.9	96.2	–	99.1
JA-31	96.2	94.8	94.8	97.9	97.9	–

[a] Nucleotide sequences of 3'-NCRs and amino acid sequences of coat protein of the 5 TuMV strains were taken from references (5, 11, 15, 17). Data of TuMV-Ca were from this study.

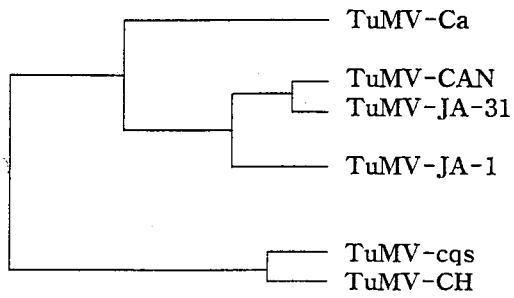

Fig. 2. Phylogenetic consensus tree of 6 TuMV strains based on the amino acid sequence alignment of their coat protein.

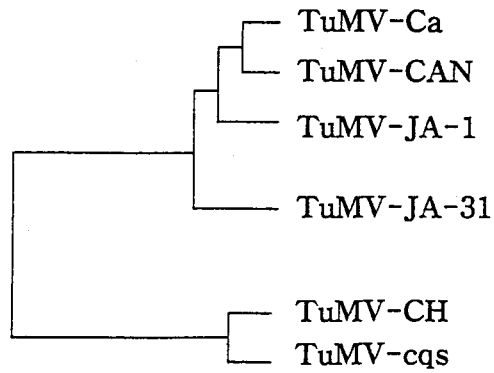

Fig. 3. Phylogenetic consensus tree of 6 TuMV strains based on the nucleotide sequence alignment of their 3'-noncoding region.

nificant identities. The 3'-NCRs of TuMV-Ca exhibits 99.5%, 94.3%, 94.7%, 99.5% and 99.5% nucleotide sequence identities to TuMV-CAN (17), TuMV-CH (11), TuMV-cqs (5), TuMV-JA-1 and TuMV-JA31 (15), respectively, showing 1 to 12 nucleotides substitutions (Table 1). On the bases of the multiple alignments of the coat proteins (CPs) and 3'-NCRs, six TuMV strains were divided into two subgroups (Fig. 2, 3). Subgroup I included TuMV-Ca, TuMV-CAN, TuMV-JA-1 and TuMV-JA31, while TuMV-cqs and TuMV-CH were grouped into subgroup II. Relationships in the 3'-NCRs of the TuMV strains were very similar with those of their CP sequence homologies. On the contrary, length of the 3'-NCRs of the known 11 potyviral RNAs and TuMV RNAs was heterogenous and their nucleotide sequences showed 55.7% to 24.0% of homology (Fig. 4). Sequence analysis showed that TuMV was more closely related to *Kalanchoë* mosaic virus (KMV) (10) than the other potyviruses (Table 2).

Sequence analysis. The 3'-NCRs of the TuMV strains did not have the general poly (A) signal sequence, AAUAAA, for poly (A) tailing (Fig. 1). Instead of this general motif, the sequence UAUGU, that has been known to be important for transcription termination in yeast (27), was found in the region from 142 to 146 bases upstream from the poly (A) tail in each TuMV strains (Fig. 1), which has been considered as another potential polyadenylational signal motif (14). Sequence analysis also revealed that a highly conserved hexanucleotide sequence, AGUGUG, found in most of other potyvirus RNAs as [AGUG A/U G/C] (4), was located in the region from 119 nucleotides from the 3'-ends in all TuMV RNAs.

Modeling of secondary structure of the 3'-NCR. The RNA secondary structures of the 3'-NCR of TuMV strains were determined by the Zuker's method of PC/GENE program and Abrahams *et al.* (2). Computer analysis of the 209 nucleotides of the 3'-NCRs of each of the TuMV strains revealed potentially 3 major stem-loop structures in the strains (Fig. 5). The structure was calculated to be −48.2 kcal. The stem-loops were denoted as I to III from the order

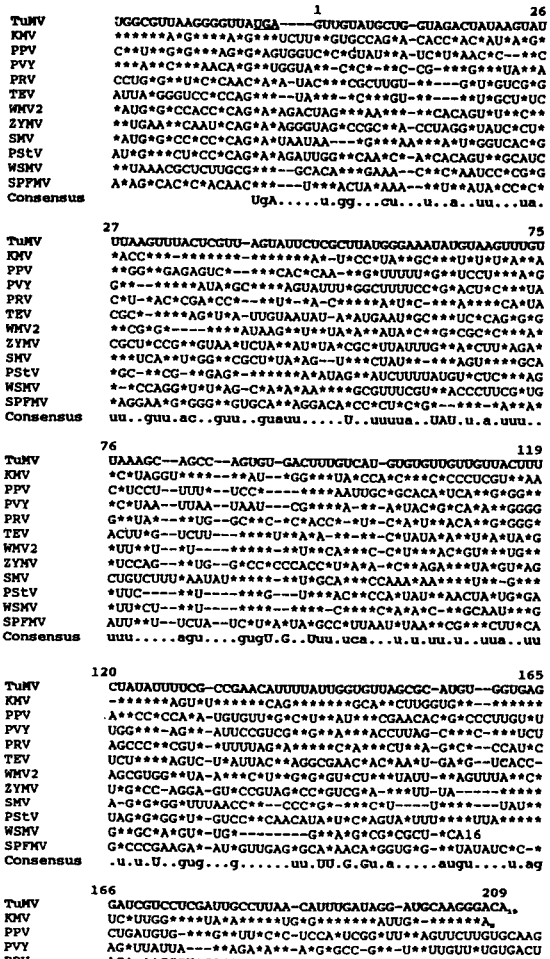

Fig. 4. Comparisons of the nucleotide sequences of the 3'-noncoding regions of TuMV and other 11 potyviruses. Nucleotides identical to the TuMV are indicated by asterisk, and gaps by dashes. Nucleotides identical in more than 6 viruses and in more than 8 viruses are shown as small and capital letters, respectively.

of the 5'-end, and their general feathers and positions are shown in Fig. 6. No alteration of the secondary structure among the six strains noted from their sub-

Table 2. Percent similarities of the 3'-noncoding region of TuMV with those of the other potyviruses

Virus	% Similarity	References
PPV	24.0	Lain *et al.* (1989)
ZYMV	31.4	Grumet and Fang (1990)
SPFMV	31.5	Abad *et al.* (1992)
PStV	37.4	Cassidy *et al.* (1993)
TEV	38.8	Allison *et al.* (1986)
PRV	38.9	Quemada *et al.* (1990)
PVY-N	39.5	Robaglia *et al.* (1989)
WSMV	41.5	Niblett *et al.* (1991)
SMV-N	47.0	Frenkel *et al.* (1989)
WMV2	49.0	Frenkel *et al.* (1989)
KMV	55.7	Husted *et al.* (1994)

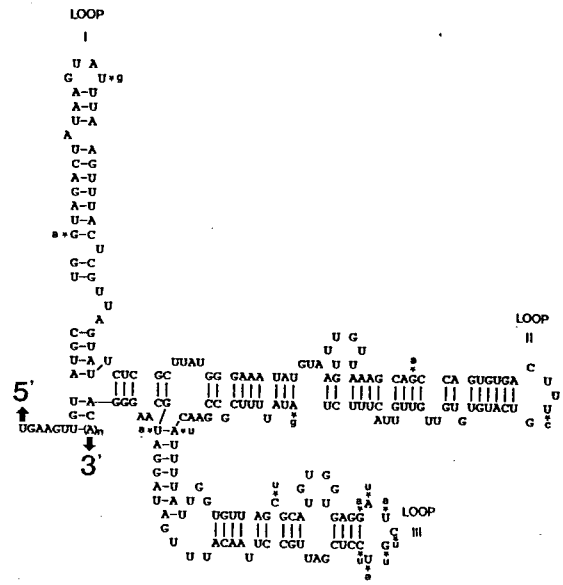

Fig. 5. Possible secondary structure of the 3'-noncoding region of TuMV. Nucleotides with asterisk indicate the substituted bases among different TuMV isolates.

stituted bases in the region. Three stem-loops were distinct in structure, especially in loop. G-U or U-G base parings were found in stem-loops I and II, and A/U bases were abundant. All loop regions were heterogenous, and their sequences were highly homologous within six TuMV isolates (Table 3).

DISCUSSION

We have cloned and determined nucleotide sequences of the 3' terminal regions for two Korean strains of TuMV (TuMV-Ca and TuMV-cqs) genomic

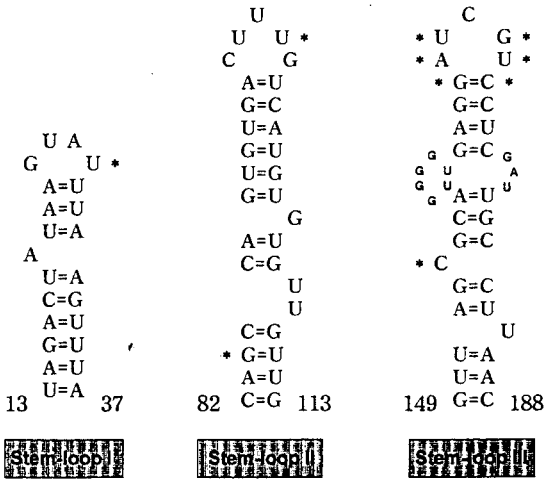

Fig. 6. Secondary stem-loop structures and their positions in the 3'-NCR of TuMV. Asterisk indicates the substituted bases among different TuMV strains.

RNAs. Our aim was to characterize TuMV strains on the bases of geographical distribution and nucleotide sequence variations such as 3'-noncoding region (3'-NCR). Most TuMV strains resemble the type strain in biological and serological properties, though some differences are found (6, 26). It is better to compare nucleotide sequences of certain viral genes or specific regions for identification and classification of TuMV strains. Nucleotide sequence of the 3'-NCRs among the TuMV strains showed significant identities from 99.5% to 94.3%. Tremblay *et al.* (24) reported nucleotide sequence of the 3'-terminal of an TuMV isolate from Canada. The length and sequence determined were not in common other strains, especially at the 3'-NCR, which was later found to be from cloning artifact by Nicolas and Laliberte (17).

Two Korean strains were homogenous with other four geographically different TuMV strains but heterogenous with other potyviruses. First, the nucleotide lengths of 3'-NCRs of the TuMV isolates were exactly the same of 209 residues long upstream of the poly (A) tail. Second, a highly conserved hexanucleotide sequence motif, AGUGUG, was found approximately 119 nucleotides from 3'-end of the NCR of TuMV strains. Third, nucleotide sequence of the 3'-ends of NCR was GAC residues and the potential polyadenylational signal motif was UAUGU. But, many members of other plant virus groups with a 3' poly (A) tail like potexviruses and carlaviruses exhibit AAUAAA for signal motif. Conclusively, TuMV is distinguished from other potyviruses. The 3'-NCR, therefore, is considered to be useful for phylogenetic studies on potyviruses. The presence of identical sequences in particular domains of TuMV and some other potyviruses implies that TuMV is originated from a common parent and were subjected to the effects of convergent evolution.

Table 3. Conserved nucleotide sequences of the loop regions in the stem-loop structures I to III located in the 3'-NCRs of six turnip mosaic virus (TuMV) strains[a]

TuMV strains	Loop nucleotide sequences		
	Loop I	Loop II	Loop III
Ca	GUAU	CUUUG	AUCGU
cqs	GUAG	CUUCG	UACUA
CH	GUAG	CUUCG	UACUA
JA-1	GUAU	CUUUG	AUCGU
JA-31	GUAU	CUUUG	AUUGU
CAN	GUAU	CUUUG	AUCGU

[a] Nucleotide sequences of 3'-NCRs of the 5 TuMV strains were taken from references (5, 11, 15, 17). Data of TuMV-Ca were from this study.

Husted *et al.* (10) reported that KMV is closely related to but distinct from TuMV based on the amino acid sequence analysis of the viral coat protein. Interestingly, our database search on the 3'-NCRs showed that TuMV is more related to KMV than other 10 potyviruses.

These information about interspecific sequence of TuMV is being used to design a PCR primer that could provide a highly sensitive and specific assay for the identification of plant tissue infected with TuMV.

요 약

순무 모자이크 바이러스(TuMV)의 한국계통인 TuMV-Ca와 TuMV-cqs 게놈 RNA의 3' 말단을 포함하는 cDNA를 사용하여 이들의 3' 비번역부위의 염기서열을 결정하였다. 두 계통 모두 3' 비번역부위는 209개의 염기로 되어 있었으며, 3' 말단은 GAC와 poly A 영역으로 구성되어 있었다. Poly A tail의 signal motif로 추정되는 UAUGU는 두 한국계통과 지금까지 보고된 4종의 외국계통 등 모든 TuMV에서 3' 말단으로부터 140번째 염기에 위치하였다. 또한 potyvirus내의 3' 비번역부위에 공통적으로 존재하는 6개의 염기로 구성된 [AGUGA/UG/C]가 TuMV에서는 3' 말단으로부터 119번째 위치에 AGUGUG로 존재하였다. TuMV 계통간 3' 비번역부위 염기서열

을 비교한 결과 한국계통은 네개의 다른 계통인 카나다, 중국 및 일본 계통에 대해 94.3~99.5%의 매우 높은 상동성을 보였다. 이들 TuMV 3' 비번역부위 염기서열을 기초로 하여 2차 구조를 분석한 결과 3개의 긴 stem-loop 구조를 이루고 있었다. 그러나 TuMV 6 계통간의 변화된 염기들은 이들 2차구조에 영향을 주지 않았다. TuMV의 3' 비번역부위를 다른 11종의 potyvirus들과 비교한 결과, 이 부분의 길이에서 차이가 있었으며, 이들 간의 염기서열 상동성은 55.7~24.0%였다. 따라서 3' 비번역부위는 potyvirus그룹의 유전적 유연관계 분석에 매우 유용하리라 생각된다.

ACKNOWLEDGEMENTS

This work was supported by grants from Plant Molecular Biology & Biotechnology Research Center (PMBBRC) in Gyeongsang National University in Chinju and the Korea Science & Engineering Foundation, Taejon, Korea.

REFERENCES

1. Abad, J. A., Conkling, M. A. and Moyer, J. W. 1992. Comparison of the capsid protein cistron from serologically distinct strains of sweet potato feathery mottle virus (SPFMV). *Arch. Virol.* 126 : 147-157.
2. Abrahams, J. P., Van den Berg, M., Batenburg, E. and Pleij, C. W. A. 1990. Prediction of RNA secondary structure, including pseudoknotting, by computer simulation. *Nucleic Acids Res.* 18 : 3035-3044.
3. Allison, R., Johnston, R. E. and Dougherty, W. G. 1986. The nucleotide sequence of the coding region of tobacco etch virus genomic RNA : evidence for the synthesis of a single protein. *Virology* 154 : 9-20.
4. Cassidy, B., Sherwood, J. L. and Nelson, R. S. 1993. Cloning of the capsid protein gene from a blotch isolate of peanut stripe virus. *Arch. Virol.* 128 : 287-297.
5. Choi, G. S. and Choi, J. K. 1993. Nucleotide sequence of coat protein gene of turnip mosaic virus (cqs strain). *Korean J. Plant Pathol.* 9 : 256-262.
6. Choi, G. S., Choi, J. K. and Lee, S. Y. 1992. Biological properties of two isolates of turnip mosaic virus isolated from Chinese cabbage. *Korean J. Plant Pathol.* 8 : 276-280.
7. Colinet, D., Kummert, J. and Lepoivre, P. 1994. The complete nucleotide sequences of the coat protein cistron and the 3' non-coding region of a newly-identified potyvirus infecting sweet potato, as compared to those of sweet potato feathery mottle virus. *Arch. Virol.* 139 : 327-336.
8. Frenkel, M. J., Ward, C. W. and Shukla, D. D. 1989. The use of 3' non-coding nucleotide sequences in the taxonomy of potyviruses : application to watermelon mosaic virus 2 and soybean mosaic virus-N. *J. Gen. Virol.* 70 : 2775-2783.
9. Grumet, R. and Fang, G. 1990. cDNA cloning and sequence analysis of the 3'-terminal region of zucchini yellow mosaic virus RNA. *J. Gen. Virol.* 71 : 1619-1622.
10. Husted, K., Bech, K., Albrechtsen, M. and Borkhardt, B. 1994. Identification, partial sequencing, and detection of a potyvirus from *Kalanchoë blossfeldiana. Phytopathology* 84 : 161-166.
11. Kong, L. J., Fang, R. X., Chen, Z. H. and Mang, K. Q. 1990. Molecular cloning and nucleotide sequence of coat protein gene of turnip mosaic virus. *Nucleic Acids Res.* 18 : 5555.
12. Lain, S., Riechmann, J. L. and Garcia, J. A. 1989. The complete nucleotide sequence of plum pox potyvirus RNA. *Virus Res.* 13 : 157-172.
13. Moghal, S. M. and Francki, R. I. B. 1976. Towards a system for the identification and classification of potyviruses. I. Serology and amino acid composition of six distinct viruses. *Virology* 73 : 350-362.
14. Mori, M., Usugi, T., Hayashi, T. and Nishiguchi, M. 1994. Nucleotide sequence at the 3'-terminal region of sweet potato feathery mottle virus (ordinary strain, SPFMV-O) RNA. *Biosci. Biotech. Biochem.* 58 : 965-967.
15. Nakashima, H., Sako, N., Joh, K., Hori, K. and Nonaka, F. 1991. Nucleotide sequence of the coat protein genes of aphid transmissible and non-transmissible isolates of turnip mosaic virus. *Ann. Phytopath. Soc. Japan* 57 : 549-557.
16. Niblett, C. L., Zagula, K. R., Calvert, L. A., Kendall, T. L., Stark, D. M., Smith, C. E., Beachy, R. N. and Lommel, S. A. 1991. cDNA cloning and nucleotide sequence of the wheat streak mosaic virus capsid protein gene. *J. Gen. Virol.* 72 : 499-504.
17. Nicolas, O. and Laliberte, J. 1991. The complete nucleotide sequence of turnip mosaic potyvirus RNA. *J. Gen. Virol.* 73 : 2785-2793.
18. Park, W. M., Choi, S. R., Ryu, K. H., Choi, C. W., Yoon, K. E. and Choi, J. K. 1993. Characterization of turnip mosaic virus isolated from Chinese cabbage. *Korean J. Plant Pathol.* 9 : 321.
19. Quemada, H., L'Hostis, B., Gonsalves, D., Reardon, I. M., Heinrikson, R., Hiebert, E. L., Sieu, L. C. and Slightom, J. L. 1990. The nucleotide sequences of the 3'-terminal regions of papaya ringspot virus strains W and P. *J. Gen. Virol.* 71 : 203-210.
20. Quemeda, H., Sieu, L. C., Siemieniak, D. R., Gon-

salves, D. and Slightom, J. L. 1990. Watermelon mosaic virus II and zucchini yellow mosaic virus : cloning of 3'-terminal regions, nucleotide sequences, and phylogenetic comparisons. *J. Gen. Virol.* 71 : 1451-1460.

21. Robaglia, C., Durand-Tardif, M., Tronchet, M., Bozadin, G., Astier-Manifacier, S. and Casse-Delbart, F. 1989. Nucleotide sequence of potato virus Y (N strain) genomic RNA. *J. Gen. Virol.* 70 : 935-947.
22. Ryu, K. H. and Park, W. M. 1994. Complementary DNA cloning and restriction mapping of nuclear inclusion body and coat protein genes of turnip mosaic virus Ca strain genomic RNA. *Korean J. Plant Pathol.* 10 : 235-239.
23. Ryu, K. H. and Park, W. M. 1994. Sequence announcement. Nucleotide sequence of partial coat protein gene for turnip mosaic virus, TuMV-Ca, the Korean isolate. *Plant Mol. Biol.* 25 : 757.
24. Tremblay, M., Nicolas, O., Shinha, R. C., Lazure, C. and Laliberte, J. 1990. Sequence of the 3'-terminal region of turnip mosaic virus RNA and the capsid protein gene. *J. Gen. Virol.* 71 : 2769-2772.
25. Van der Vlugt, R. A. A., Leunissen, J. and Goldbach, R. 1993. Taxonomic relationships between distinct potato virus Y isolates based on detailed comparisons of the viral coat protein and 3'-nontranslated regions. *Arch. Virol.* 131 : 361-375.
26. Yoshii, H. 1963. On the strain distribution of turnip mosaic virus. *Ann. Phytopath. Soc. Japan* 28 : 221-227.
27. Zaret, K. S. and Sherman, F. 1982. DNA sequences required for efficient transcription termination in yeast. *Cell* 28 : 563-573.