Nucleotide Sequence of the Coat Protein Gene and 3'-noncoding Region of a Japanese Isolate of WMV 2*

Keiko YOSHIOKA**, Kaoru HANADA***, Chiyoichi NODA† and Katsuji OOSAWA††

Key words : watermelon mosaic virus 2, coat protein gene, 3'-noncoding region.

Potyvirus is a large and economically important group of plant viruses. Watermelon mosaic virus 2 (WMV 2), which belongs to the genus *Potyvirus*, is one of the serious pathogens in Cucurbitaceae. The genome of WMV 2 consists of a single-stranded, plus-sense RNA, and has a poly(A) tail at the 3' end.

We determined the nucleotide sequence of the 3'-terminal region containing the complete coding region of the coat protein gene of a Japanese isolate of WMV 2.

The Japanese isolate of WMV 2 (WMV 2-80) used in this work was described by Yamamoto *et al.*⁸⁾ The virus was propagated in cucumber and purified according to the procedure previously described⁵⁾. Viral RNA was extracted by the method described by Quemada *et al.*⁴⁾ cDNAs were synthesized according to the method of Gubler and Hoffman³⁾ using oligo dT primers and ligated into the *Pst*I site of pBR 322. These cDNAs were subcloned into the *Pst*I site of pBluescript II SK+ and different-sized cDNA clones were obtained. Among them, cDNA clones W17 and W46, both of which had inserts of approximately 1.7kb, were used for determination of nucleotide sequence. Nucleotide sequences of deletion clones obtained from the two cDNAs were determined by the dideoxy-chain termination method⁶⁾ using the M13, M4 and RV primers with a DNA sequencer model 373A (Applied Bio Systems Co., Ltd.). Computer analyses of nucleotide sequences and amino acid sequences were done by the DNASIS computer program (Hitachi Software Engineering Co., Ltd.). There was no difference in nucleotide sequences between the overlapping regions of the W17 clone and the W46 clone.

The nucleotide and deduced amino acid sequences of the cDNA clone are shown in Fig. 1. Nucleotide sequence comparisons of this sequence with other isolates already reported^{2,4,7)} indicate that this cDNA insert contains the entire coat protein gene and the 3' untranslated region. The region encoding the coat protein of the Japanese isolate is 849 nucleotides in length and deduced coat protein is 283 amino acids. The coat protein of the

Japanese isolate contains a stretch of two amino acids that are absent from sequences of the other isolates. These two amino acids are inserted in the position immediately downstream of the proposed cleavage site recognized by the NIa proteinase (Fig. 1). The 3' untranslated region was 243 nucleotides long and included the possible yeast polyadenylation signal TATGT¹¹⁾ at 64 nucleotides downstream from the stop codon of the coat protein coding region (Fig. 1). This signal has been observed in the other isolates of WMV 2. However the signal position of the Japanese isolate is different from those of the U.S.A. and Australian isolates which are 115 and 111 nucleotides downstream from the stop codon of the coding sequence of the coat protein, respectively.

The amino acid sequence near the proposed cleavage site of the WMV 2 Japanese isolate is aligned with those of other isolates of WMV 2^{2,4,7)}, U.S.A. isolate of zucchini yellow mosaic virus⁴⁾, U.S.A. isolate of soybean mosaic virus (SMV)¹⁰⁾, and U.S.A. isolate of bean yellow mosaic virus⁹⁾ (Table 1). Identical amino acids are indicated by a box and stippled boxes. The sequence of inserted amino acids of the Japanese isolate was found in the same position of the U.S.A. isolate of SMV, but not in any other WMV 2 isolates sequenced so far (Table 1). Homology of nucleotide sequence and amino acid sequence found between Japanese isolate and other isolates is shown in Table 2. The nucleotide sequence of the coat protein gene of Japanese isolate shares 92, 93 and 92% identity with those of U.S.A., Australian and Tongan isolates, respectively. Its deduced amino acid sequence shares 93, 95 and 94% identity. The major difference of amino acid sequence among the WMV 2 isolates was found in the N-terminus, whereas the middle core region and C-terminus was highly conserved apart from the insertion. The Japanese isolate was closely related to the Australian, Tongan and the U.S.A. isolate with regard to the coat protein coding region. The nucleotide sequence of the 3' untranslated region of the Japanese isolate was more closely related to the

* The nucleotide sequence data reported here will appear in the DDBJ /EMBL/GenBank nucleotide sequence databases under the accession number D85837.

** The Institute of Physical and Chemical Research (RIKEN), Wako 351-01, Japan 理化学研究所

*** Kyushu National Agricultural Experiment Station, Kikuchi-gun, Kumamoto 861-11, Japan 九州農業試験場

† Japan International Research Center for Agriculture Sciences, Ishigaki 907, Japan 国際農業研究センター沖縄支所

†† National Institute of Agrobiological Resources, Tsukuba 305, Japan 農業生物資源研究所

1	-CTTCATTTGCTGAGCTTGGTCTGAATTATAATTTTGTATGAGCGAACAAAGAAAAGAGAG	60
1	X S F A E L G L N Y N F D E R T K K R E	20
61	GAACCTGTGGTTTCATGTACACCAAGCCATGCTAGTTGATGGGATATACATCCCGAAGCTT	120
21	E L W F M S H Q A M L V D G I Y I P K L	40
121	GAACCTGAAAGATTAGTTTCAATCTTAGAGTGGGATAGGAGCAAAGAAATTATGCACCGC	180
41	E P E R L V S I L E W D R S K E I M H R	60
181	ACTGAGGCAATTTGTGCGCCATGATTGAAGCATGGGGATATACTGAGCTTTTGCAAGAA	240
61	T E A I C A A M I E A W G Y T E L L Q E	80
241	ATTCGCAAATTTTATTGTGGCTTCTAAGCAAAGATGAATTTAAGGAACTCGCTGCATCT	300
81	I R K F Y L W L L S K D E F K E L A A S	100
301	GGAAAAGCTCCATACATAGCTGAGACAGCATTGAGAAAGCTCTATACAGACGTCAACACA	360
101	G K A P Y I A E T A L R K L Y T D V N T	120
361	CAACCAAGTGAATTACAAAAATACCTTGAAGTTCTGGATTTCACCATACTGATGGTTGC	420
121	Q P S E L Q K Y L E V L D F N H T D G C	140
421	TGCGAATCAGTGTCTCTGCAATCAGGAAAAGAAAAGAAACAGTGGAAAATTTGGATGCA	480
141	C E S V S L Q S G K E T V E N L D A	160
481	GGAAAGGACTCAAAGAAAGACACCAGTGGTAAAGGTGATAAGCCGCAAAACTCACAACT	540
161	G K D S K K D T S G K G D K P Q N S Q T	180
541	GGGCAGGGTAGCAAGGAACAGCAAAAGCTGGCACAGTCAGCAAAGATGTGGACGCTGGG	600
181	G Q G S K E P A K A G T V S K D V D A G	200
601	TCAAAAGGAAAAGTAGTCCCACGACTACAAAAGATAACAAGAAGATGAACCTTCCAACA	660
201	S K G K V V P R L Q K I T K K M N L P T	220
661	GTTGGTGGAAAAATCATTCTTAGCTTAGACCATTGCTCGAGTACAAACCGAATCCAGTC	720
221	V G G K I I L S L D H L L E Y K P N P V	240
721	GATCTGTTCAACACTCGAGCAACAAAGACACAGTTTGAATCATGGTATAGCGCAGTTAAA	780
241	D L F N T R A T K T Q F E S W Y S A V K	260
781	ACTGAATATGATCTTAATGATGAGCAAATGGGTGTGATTATGAATGGTTTATGGTTTGG	840
261	T E Y D L N D E Q M G V I M N G F M V W	280
841	TGCATCGATAATGGTACATCTCCAGATGTTAATGGAGTATGGGTGATGATGGATGGGAA	900
281	C I D N G T S P D V N G V W V M M D G E	300
901	GAGCAAGTTGAGTACCCACTAAAGCCAATTGTTGAAAATGCAAAGCCAACCTTTGAGACAA	960
301	E Q V E Y P L K P I V E N A K P T L R Q	320
961	ATCATGCACCATTTCTCAGACGAGCAGAAGCATATATTGAAATGAGAACTCTGAAAGT	1020
321	I M H H F S D A A E A Y I E M R N S E S	340
1021	CCGTATATGCCTAGATACGGATTACTAAGAAATTTGAGAGACAGGGAATTAGCACGCTAT	1080
341	P Y M P R Y G L L R N L R D R E L A R Y	360
1081	GCTTTTGACTTTTATGAGGTTACTTCCAAAACACCAAATAGGGCAAGAGAAGCAATAGCA	1140
361	A F D F Y E V T S K T P N R A R E A I A	380
1141	CAAATGAAGGCCGAGCTCTCGCGGAATTAACAGCAGGTTATTTGGGCTTGATGGTAAT	1200
381	Q M K A A A L A G I N S R L F G L D G N	400
1201	ATCTCGACCAATTCGAAAATACTGAGAGGCACACTGCAAGGGACGTGAATCAGAACATG	1260
401	I S T N S E N T E R H T A R D V N Q N M	420
1261	CATACTTTGTTGGGTATGGGTCCACCGCAGTAAAGACTAGGTAAACTGGTCACAGTTAGC	1320
421	H T L L G M G P P Q *	431
1321	ATTTCCGGTTCGTTATAAGTTTCTATAATAATAATATGTTGCACTTTATTTTAGTATAGTG	1380
1381	TGATTTTCATCACCTTTGTACTGTTTATGTTAGCCTGGTTTAACCACCTTAGTGTGTGCTT	1440
1441	TATATTATAGTTTATGCGTAGCAGGGAGAACCATTACAATGCCGGAGTTGTTTGTAGTGT	1500
1501	GATTTTCATCACGGTTAATAGCCGAGGTACGGTAATGTTTGTTCCTAAAAA	1560
1561	AAAAAAA	

Fig. 1. Nucleotide and deduced amino acid sequences of 3' terminal regions of WMV 2 Japanese isolate. An arrowhead indicates the cleavage site recognized by proteinase. An asterisk indicates the position of termination codon. The possible yeast polyadenylation signal is underlined. The hatched amino acids are insertion specific to the Japanese isolate among WMV 2 (see Table 1).

U.S.A. isolate than to the Australian and Tongan isolates. This is the first report of the nucleotide sequence of 3'

terminal region containing the complete coding region of the coat protein gene of a Japanese isolate of WMV

Table 1. Amino acid sequence comparison at the proposed cleavage sites recognized by the proteinase

Virus	Isolate	Amino acid position ^{a)}														
		-4	-3	-2	-1	↓ ^{b)}	1	2	3	4	5	6	7	8	9	10
WMV 2	JAPAN	V	S	L	Q	/	S	G	K	E	K	E	T	V	E	N
WMV 2	AUST.	V	S	L	Q	/	S	G	K	E	A	V	E	N	L	D
WMV 2	U.S.A.	V	M	L	Q	/	S	G	K	E	T	V	E	N	L	D
WMV 2	Tonga	V	S	L	Q	/	S	G	K	E	T	V	E	N	L	D
ZYMV	U.S.A.	V	M	L	Q	/	S	G	T	Q	P	T	V	A	D	A
SMV	U.S.A.	V	S	L	Q	/	S	G	K	E	K	E	Q	D	M	D
BYMV	U.S.A.	C	R	F	Q	/	S	D	Q	E	Q	L	N	A	G	E

Identical amino acids are indicated by stippled boxes and a box. The inserted amino acids are indicated by underline.

a) Position is the number from the cleavage site.

b) The cleavage site recognized by the proteinase.

Table 2. Nucleotide and amino acid sequence homology (%) between WMV 2 Japanese isolate and other isolates^{a)}

Isolate	Coat protein		3'un-translated region
	Nucleotide	Amino acid	Nucleotide
U.S.A.	92	93	98
Australia	93	95	92
Tonga	92	94	96

a) Homology was calculated based on the DNASIS computer program (Hitachi Software Engineering Co., Ltd.).

2. Recently, many studies on development of virus resistant plants by transformation with the coat protein gene were reported¹⁾.

We have embarked on this study to breed WMV 2 resistant crops expressing the coat protein gene of the Japanese isolate and have already produced some transgenic melon plants. Analysis of the coat protein mediated resistance to the virus in transgenic melon plants is now in progress.

Literature cited

1. Beachy, R.N., Loesch-Fries and Tumer, N.E. (1990). Coat protein-mediated resistance virus infection. *Annu. Rev. Phytopathol.* 28 : 451-474.
2. Frankel, 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 water melon mosaic virus 2 and soybean mosaic virus-N. *J. Gen. Virol.* 68 : 297-304.
3. Gubler, U. and Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* 25 : 263-269.
4. Quemada, H., Sieu, L.C., Siemieniak, D.R., Gonsalves, 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.
5. Sako, N., Matsuo, K. and Nonaka, F. (1980). Purifica-

tion of water melon mosaic virus. *Ann. Phytopathol. Soc. Jpn.* 46 : 639-646.

6. Sanger, F., Coulson, A.R., Barrell, B.G., Smith, A.J.H. and Roe, B.A. (1980). Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. *J. Mol. Biol.* 143 : 161-178.
7. Wang, Y.Y., Beck, D.L., Gardner, R.C. and Person, M.N. (1993). Nucleotide sequence, serology and symptomatology suggest that vanilla necrosis potyvirus is a strain of watermelon mosaic virus II. *Arch. Virol.* 129 : 93-103.
8. Yamamoto, T., Ishii, M. and Katube, T. (1982). Serological grouping of watermelon mosaic virus isolates. *Ann. Phytopathol. Soc. Jpn.* 48 : 613-619.
9. Hammond, J. and Hammond, R.W. (1989). Molecular cloning, sequencing and expression in *Escherichia coli* of the bean yellow mosaic virus coat protein gene. *J. Gen. Virol.* 70 : 1961-1974.
10. Yu, M.H., Frenkel, M.J., McKern, N.M., Shukla, D.D., Strike, P.M. and War, C.W. (1989). Coat protein of potyviruses. 6. Amino acid sequences suggest water melon mosaic virus 2 and soybean mosaic virus-N are strains of the same potyvirus. *Arch. Virol.* 105 : 55-64.
11. Zaret, K.S. and Sherman, E. (1982). DNA sequence required for efficient transcription termination in yeast. *Cell* 28 : 563-535.

和 文 摘 要

吉岡啓子・花田 薫・野田千代一・大澤勝次：日本で分離したカボチャモザイクウイルス (WMV 2) ゲノムの外被タンパク質 (CP) 遺伝子とその 3' 非翻訳領域の塩基配列の決定

日本で分離した WMV 2 の CP 遺伝子とその 3' 非翻訳領域の塩基配列を決定した。CP 遺伝子の領域は、849 塩基数でアミノ酸は 283 残基であった。このアミノ酸数は既報の豪州、米国、トンガ分離株よりも 2 残基多いものであった。この 2 残基は、翻訳後の CP 切断部位の近傍にあり、反復配列であった。同じ配列が、ダイズモザイクウイルスの米国防分離株にみられ、この株との類縁性が示唆された。塩基配列を上記の分離株と比較したところ、CP 遺伝子部分で約 92~93 %、3' 非翻訳領域で 92~98 % の相同性であった。

(Received December 4, 1995; Accepted April 30, 1996)