NUCLEOTIDE SEQUENCES OF THE COAT PROTEIN GENES OF TWO JAPANESE ZUCCHINI YELLOW MOSAIC VIRUS ISOLATES

A.K. KUNDU, K.OHSHIMA*, N. SAKO

Laboratory of Plant Virology, Faculty of Agriculture, Saga University, 1-banchi, Honjo-machi, Saga 840, Japan

Received July 11, 1997

Summary. – The nucleotide (nt) sequences of the coat protein (CP) genes of two Japanese zucchini yellow mosaic virus (ZYMV) isolates (ZYMV-169 and ZYMV-M) were determined. The CP genes of both isolates were 837 nt long and encoded 279 amino acids (aa). The nt and deduced aa sequence similarities between the two isolates were 92% and 94.6%, respectively. The deduced aa sequences of CPs of the Japanese isolates were compared with those of previously reported ZYMV isolates by phylogenetic analysis. This comparison lead us to divide all ZMYV isolates into 3 groups in which ZYMV-169 formed its own distinct group.

Key words: zucchini yellow mosaic virus; Japanese isolates; coat protein gene; phylogenetic analysis

ZYMV (genus *Potyvirus*, family *Potyviridae*) was first reported in Italy in 1973 (Lisa *et al.*, 1981). ZYMV has been associated in the last decade with severe economical losses in cucurbit crops in many parts of the world (Lisa and Lecoq, 1984). Potyviruses including ZYMV contain a genome that is translated as a large polyprotein precursor which is subsequently cleaved by virus-coded proteolytic enzymes to eight proteins (Dougherty and Carrington, 1988). The CP is located at the C-terminus of the polyprotein (Dougherty and Carrington, 1988). The use of CP aa sequences has recently become a useful tool for taxonomic classification of distinct potyviruses and related strains (Shukla *et al.*, 1994).

The nt sequences of CP genes of ZYMV isolates from the USA (Grumet and Fang, 1990; Quemada *et al.*, 1990; Balint *et al.*, 1994), Israel (Gal-On *et al.*, 1990), Singapore (Lee *et al.*, 1993) and Reunion Island (Baker *et al.*, 1994) have been reported.

In Japan, ZYMV was first isolated from diseased pumpkin from Okinawa island by Ohtsu *et al.* (1985) and was designated as ZYMV-169. Recently, we obtained another isolate (ZYMV-M) from commercial cucumber field in Miyazaki prefecture, Kyushu island. These two isolates did not differ in host range properties, but showed some serological differences using polyclonal antibodies (PAbs) (Kundu *et al.*, 1996). Although ZYMV has been recorded in Japan since 1985, no molecular analysis of its genome sequence variability has been reported. In this paper, we determined the nt sequences of CP genes of these two isolates and described the phylogenetic relationships among the Japanese and previously reported ZYMV isolates.

The Japanese ZYMV isolates were propagated on *Cucurbita maxima* Duch. cv. Hokoseihi and purified according to Sako *et al.* (1980). The genomic RNAs of ZYMV isolates were extracted by the procedure described by Rosner *et al.* (1983). The method of reverse transcription and polymerase chain reaction (RT-PCR) of Ohshima *et al.* (1994) was employed for cDNA cloning of the CP genes with some modifications. First-strand cDNAs were synthesized from ZYMV RNAs with minus strand oligonucleotide primer ZY3-1M (5'-GGGGAATTCCG-GAGTCTAATCTCGAGC-3') using MMLV reverse transcriptase (Gibco BRL). Double-stranded cDNAs were amplified from the first-strand cDNAs using oligonucleotide

*Corresponding author.

Abbreviations: aa = amino acid; CP = coat protein; MMLV = Moloney murine leukemia virus; NIb = nuclear inclusion b; nt = nucleotide; PAb = polyclonal antibody; RT-PCR = reverse transcription-polymerase chain reaction; SDS = sodium dodecyl sulfate; 3'UTR = 3'-untranslated region; ZYMV = zucchini yellow mosaic virus

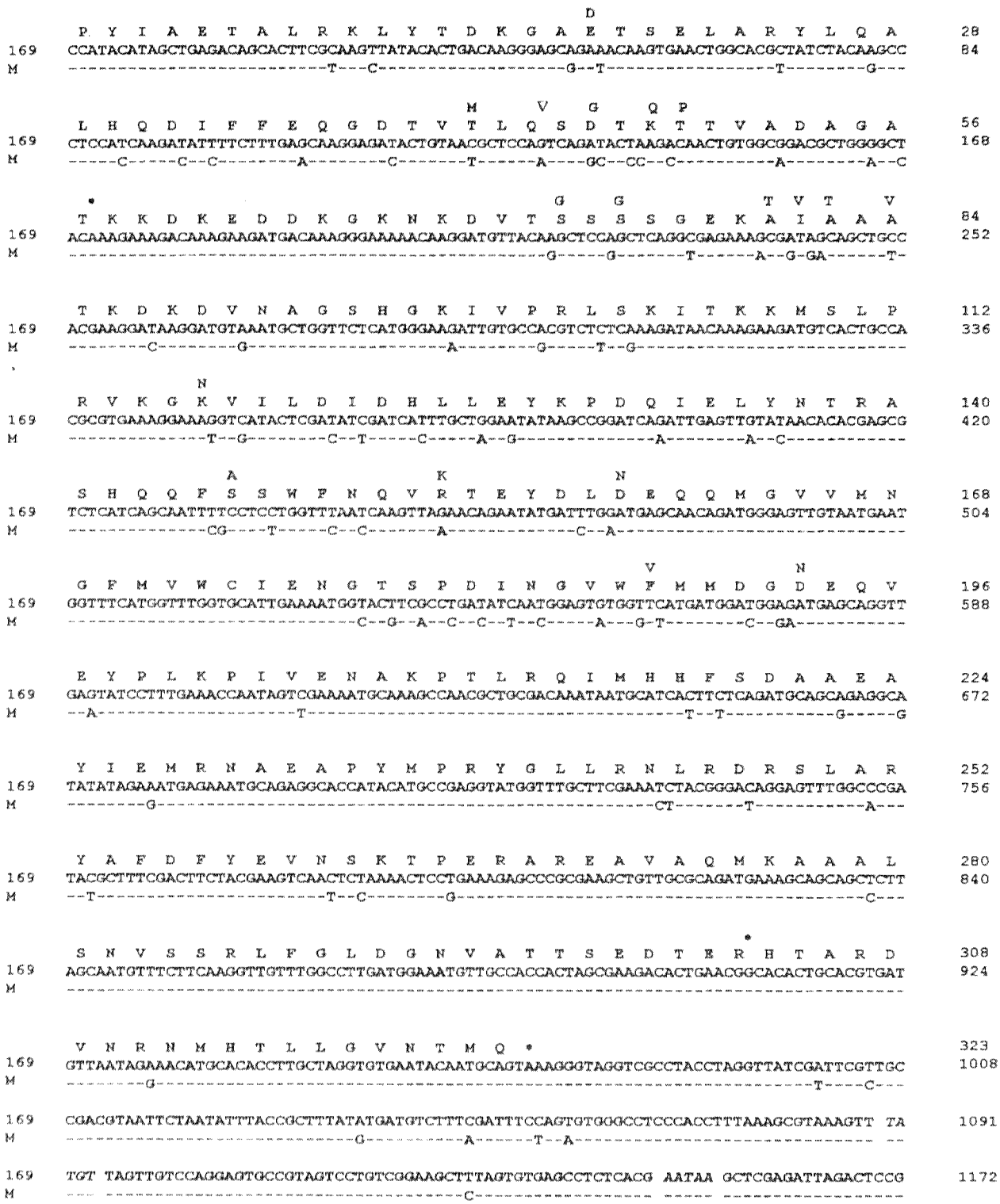

Fig. 1

Comparison of nt and deduced aa sequences of 3'-terminal regions of ZYMV-169 and ZYMV-M cDNAs

Dashes show identical nt in comparison to ZYMV-169. The deduced aa sequence of ZYMV-169 is shown above its nt sequence. The different aa found in ZYMV-M are shown above the aa of ZYMV-169. The putative protease cleavage site between NIb and CP is denoted by triangle (V). Approximate N-and C-terminal regions as suggested by Shukla *et al.* (1994) are denoted by black dot (●). Asterisk (*) indicates the ochre codon. Italics show potential polyadenylation signals.

```
          10         20         30         40  •      50         60         70         80
169 SDTKTTVADA GATKKDKEDD KGKNKDVTSS SSGEKAIAAA TKDKDVNAGS HGKIVPRLSK ITKKMSLPRV KGKVILDIDH
M   -G-QP----- ---------- --------G- G----TVT-V ---------- ---------- ---------- --N-------
Cal -G-QP----- ---------- --------G- G----TV--V ---------- ---------- ---------- --N-------
Con -G-QP--S-- ---------- --------G- G----TV--V ---------- ---------- ---------- --N-------
F   -G-QP----- RV-------- --E---F-G- G----TVV-- K--------- ---------- ---------- --N-------
NAT -G-QP----T ---------- --------G- G-S--TV--V ---------- ---------- ---------- --N-------
R   ---Q-KE--- --A-R--DEE -E-K---A-- -AN--TMT-T A--------- ---------- ---------- --S-------
S   ---Q-RE-G- --S----DE- -D-K---A-- -AS---V-T- ---------- ---------- ---------- --S-------

          90        100        110        120        130        140        150        160
169 LLEYKPDQIE LYNTRASHQQ FSSWFNQVRT EYDLDEQQMG VVMNGFMVWC IENGTSPDIN GVWFMMDGDE QVEYPLKPIV
M   ---------- ---------- -A------K- ----N----- ---------- ---------- ---V----N- ----------
Cal ---------- ---------- -A------K- ----N----- ---------- ---------- ---V----N- ----------
Con ---------- ---------- -A------K- ----N----- ---------- ---------- ---V----N- ----------
F   ---------- ---------- -A------K- ----ND---- ---------- ---------- --------N- ----------
NAT ---------- ---------- -A------K- ----N----- ---------- ---------- ---V----N- ----------
R   ---------- ---------- -A------KA ----N----- ---------- ---------- ---------- ----------
S   ---------- ---------- -A------KA ----N----- ---------- ---------- ---V----N- ----------

          170        180        190        200        210        220        230        240
169 ENAKPTLRQI MHHFSDAAEA YIEMRNAEAP YMPRYGLLRN LRDRSLARYA FDFYEVNSKT PERAREAVAQ MKAAALSNVS
M   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Cal ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
Con ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
F   ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
NAT ---------- ---------- ---------- ---------- ---------- ---------- ---------- ----------
R   ---------- ---------- ---------- ---------- ---------- ---------- -D-------- ----------
S   ---------- ---------- ---------- ---------- ---------- ---------- -D-------- ----------

          250                   270       279
169 SRLFGLDGNV ATTSEDTERH TARDVNRNMH TLLGVNTMQ
M   ---------- ---------- ---------- ---------
Cal ---------- ---------- ---------- ---------
Con ---------- ---------- ---------- ---------
F   ---------- ---------- ---------- ---------
NAT ---------- ---------- ---------- ---------
R   ---------- --N------- ---------- ---------
S   ---------- ---------- ---------- ---------
```

Fig. 2

Alignment of deduced aa sequences of CPs of ZYMV isolates

Dashes show identical aa in comparison to ZYMV-169. Approximate N- and C-terminal regions as suggested by Shukla *et al.* (1994) are denoted by black dot (●). Sources of the CP sequences are as follows; ZYMV-Cal (Balint *et al.*, 1994), ZYMV-Con (Grumet and Fang, 1990), ZYMV-F (Quemada *et al.*, 1990), ZYMV-NAT (Gal-On *et al.*, 1990), ZYMV-R (Baker *et al.*, 1994) and ZYMV-S (Lee *et al.*, 1993).

primer pairs W.Z.NIb-1P (5'-GGGGGATCCCATACAT-AGCTGAGACAGC-3') and ZY3-1M by the PCR method (Saiki *et al.*, 1988). The amplified double-stranded cDNAs were cloned into pBluescript II SK$^+$ vector (Stratagene) and then the obtained recombinant plasmids were introduced into *Escherichia coli* XL1-Blue. The DNAs were sequenced using ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit. The nt and deduced aa sequence analyses, similarity searches and alignments were carried out using the SDC-GENETYX computer program (Software Development). Phylogenetic analyses of aa sequences of CPs and nt sequences of 3'-untranslated regions (3'UTRs) were accomplished using the PROTPARS and DNAPARS programs of PHYLIP (Phylogeny Inference Package), respectively, as developed by Felsenstein (1993).

For expression of CPs in *E. coli*, three oligonucleotide primers, namely ZYMV-169CPN (5'-GGGGGATCC TCTCAGA-TACTAAGACAACTGTGGCGGACGCTG-3'), ZYMV-MCPN (5'-GGGGGATCCTCTCAGGCACCCA GCCAACTGTGGCAGACGCTG-3') and ZYMVCP (5'-GGGGAATTCTCACTGCATTGTATTCACACCTAG-3') were designed on the basis of the nt sequences of ZYMV-169 and ZYMV-M determined in this study. First-strand cDNAs of ZYMV-169 and ZYMV-M were synthesized from respec-

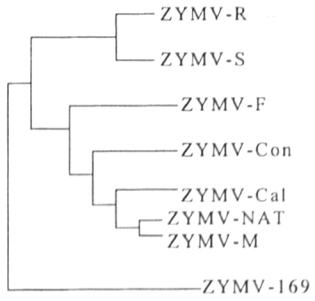

Fig. 3
Phylogenetic analysis of ZYMV isolates based on aa sequences of their CPs
The dendrogram was drawn by using the data from Fig. 2 and the PROTPARS (PHYLIP) program.

tive RNA using the minus strand oligonucleotide primer ZYMVCP. Double-stranded cDNAs were amplified from ZYMV-169 and ZYMV-M cDNAs using oligonucleotide primer pairs ZYMV-169CPN and ZYMVCP, and ZYMV-MCPN and ZYMVCP, respectively, by the PCR method (Saiki *et al.*, 1988). The amplified double-stranded cDNAs were cloned into pGEX3X vector (Pharmacia) and then the recombinant plasmids were introduced into *E. coli* XL1-Blue. Expression of the recombinant plasmids was conducted as described by Ohshima *et al.* (1994). The expressed proteins were electrophoresed in 10% polyacrylamide gels in the presence of sodium dodecyl sulfate (SDS) as described by Laemmli (1970). Western blot analysis was performed according to Towbin *et al.* (1979) using PAbs against ZYMV-169 and ZYMV-M, respectively, and horseradish peroxidase-labelled goat anti-rabbit immunoglobulin G (Zymed).

More than three recombinant clones for each isolate were obtained and their nt sequences were determined in both orientations. No differences were observed in overlapping regions of these clones. In order to verify the cloned CP genes of both isolates they were recloned into pGEX3X and expressed in *E. coli* as fusion proteins with glutathione S-transferase. The fused proteins positively reacted with PAbs against ZYMV-169 and ZYMV-M, respectively, in Western blot analysis (data not shown).

Nucleotide sequences of the cDNAs including a portion of NIb gene, the complete CP gene and 3'UTR (without poly(A)-tail) with deduced aa sequences are shown in Fig. 1. The nt sequences of CP genes of ZYMV-169 and ZYMV-M were 837 nt long and encoded 279 aa. A DAG sequence essential for aphid transmission (Atreya *et al.*, 1990) was present at the N-terminal region of CP of both isolates. Furthermore, possible eukaryotic polyadenylation signals AATAA and TATGT (Zaret and Shermann, 1982) were found at 3'UTR of both isolates. Comparison of the nt and deduced aa sequences of CP genes of ZYMV-169 and ZYMV-M revealed 92% and 94.6% similarity, respectively. Seventy nt substitutions were present between the CP genes of the two isolates and gave rise to 15 aa differences. Out of the 15 aa differences, 9 were located at the N-terminal region and the remaining 6 were at the core region (Fig. 1). It is known that the N-terminal region of CP of potyviruses is immunodominant, involved in virus-specific immunochemical properties (Shukla *et al.*, 1994).

The deduced aa sequences of CP genes of Japanese ZYMV isolates were compared with those of previously reported ZYMV isolates (Fig. 2). ZYMV-169 showed 91% – 94.6% aa similarity to other ZYMV isolates and had unique aa at positions 4, 37, 73, 102, 109 and 115. On the other hand, ZYMV-M showed 91% – 99.6% aa similarity and did not contain a unique aa. A phylogenetic analysis derived from the CP aa sequences revealed 3 groups among the 8 ZYMV isolates examined (Fig. 3). One group contained ZYMV-R and ZYMV-S, and another one contained ZYMV-F, ZYMV-Con, ZYMV-Cal, ZYMV-NAT and ZYMV-M. On the other hand, ZYMV-169 formed its own group. Furthermore, a phylogenetic analysis of the nt sequences from 3'UTRs of the ZYMV isolates was performed. The dendrogram of the 3'UTRs showed a pattern similar to that of the CPs. Therefore, ZYMV-169 isolated from Okinawa island is considered a distinct isolate differing from the others.

Acknowledgement. We thank the Structure-Function Analysis System for Useful Biological Macromolecules of Saga University for technical help.

Footnote. The nucleotide sequence data of ZYMV-169 and ZYMV-M reported in this paper will appear in the DDBJ, EMBL and GenBank nucleotide sequence data bases under accession No. AB004640 and AB004641, respectively.

References

Atreya CD, Raccah B, Pirone TP (1990): A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. *Virology* **178**, 161–165.

Baker CA, Marlow GC, Wisler GW, Lecoq H, Hiebert E (1994): Comparative sequence analysis of zucchini yellow mosaic potyvirus (ZYMV) from Reunion Island with other ZYMV isolates. *GenBank Accession No.* **L31350**.

Balint R, Plooy I, Steele C (1994): The nucleotide sequence of zucchini yellow mosaic potyvirus. *GenBank Accession No.* **L29569**.

Dougherty WG, Carrington JC (1988): Expression and function of potyviral gene products. *Annu. Rev. Phytopathol.* **26**, 123–143.

Felsenstein J (1993): *PHYLIP (Phylogeny Inference Package), Version 3.5c.* University of Washington, Seattle.

Gal-On A, Antignus Y, Rosner A, Raccah B (1990): Nucleotide sequence of the zucchini yellow mosaic virus

capsid-encoding gene and its expression in *Escherichia coli*. *Gene* **87**, 273–277.

Grumet R, 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.

Kundu AK, Ohshima K, Sako N (1996): Nucleotide sequences of the coat protein genes of two zucchini yellow mosaic virus isolates. *Ann. Phytopathol. Soc. Japan* **62**, 335 (in Japanese).

Laemmli UK (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.

Lee SC, Wu M, Wong SM (1993): Nucleotide sequence of a Singapore isolate of zucchini yellow mosaic virus coat protein gene revealed an altered DAG motif. *Virus Genes* **7**, 381--387.

Lisa V, Boccardo G, D'Agostino G, Dellavalle G, D'Aquilla M (1981): Characterization of a potyvirus that cause zucchini yellow mosaic virus. *Phytopathology* **71**, 667–672.

Lisa V, Lecoq H (1984): Zucchini yellow mosaic virus. *CMI/AAB Descriptions of Plant Viruses*, No. **282.**

Ohshima K, Matsuo K, Sako N (1994): Nucleotide sequences and expression in *Escherichia coli* of the coat protein genes from two strains of melon necrotic spot virus. *Arch. Virol.* **138**, 149–160.

Ohtsu Y, Sako N, Somowiyarjo S (1985): Zucchini yellow mosaic virus isolated from pumpkin in Miyako and Yaeyama islands, Okinawa, Japan. *Ann. Phytopathol. Soc. Japan* **51**, 234–237 (in Japanese).

Quemada H, Seiu LC, Siemieniak DR, Gonsalves D, Slighton JL (1990): Watermelon mosaic virus 2 and zucchini yellow mosaic virus: cloning of 3'-terminal regions, nucleotide sequences and phylogenetic relationships. *J. Gen. Virol.* **71**, 1451–1460.

Rosner A, Ginzburg I, Bar-Joseph M (1983): Molecular cloning of complementary DNA sequences of citrus tristeza virus RNA. *J. Gen. Virol.* **64**, 1757–1763.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988): Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* **230**, 1350–1354.

Sako N, Matsuo K, Nonaka F (1980): Purification of watermelon mosaic virus. *Ann. Phytopathol. Soc. Japan* **46**, 639–646 (in Japanese).

Shukla DD, Ward CW, Brunt AA (1994): *The Potyviridae*. CAB International, Wallingford, pp. 113-208.

Towbin H, Staehelin T, Gordon J (1979): Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and applications. *Proc. Natl. Acad. Sci. USA* **76**, 4350–4354.

Zaret KS, Shermann F (1982): DNA sequence required for efficient transcription termination in yeasts. *Cell* **28**, 563–573.