The complete nucleotide sequence of turnip mosaic virus 
RNA Japanese strain*

Brief Report

K. Ohshima, M. Tanaka, and N. Sako

Laboratory of Plant Virology, Faculty of Agriculture, Saga University, Saga, Japan

Accepted June 17, 1996

Summary. The complete nucleotide sequence of the RNA genome of turnip 
mosaic virus Japanese strain (TuMV-J) has been determined from five overlap-
ning cDNA clones and by direct sequencing of viral RNA. The RNA sequence 
was 9,833 nucleotides in length, excluding a 3' terminal poly(A) tail. An AUG 
triplet at position 130–132 was assigned as the initiation codon for the trans-
lation of the genome size viral polyprotein which would consist of 3,164 amino acid 
residues. Interestingly, a different amino acid sequence (continuous twenty 
amino acids) within the cytoplasmic inclusion protein between TuMV-J and 
Canadian strain of TuMV was observed, caused by an insertion and a deletion of 
nucleotides.

Turnip mosaic virus (TuMV), a member of the genus Potyvirus in the family 
Potyviridae, causes diseases worldwide [25]. The monopartite genome of 
potyviruses consists of a single-stranded, positive-sense approximately 10 kb 
RNA molecule containing a single open reading frame whose translation 
product is processed by cis- and trans-activating viral proteinases to yield mature 
viral proteins [2]. To date, the complete nucleotide sequences of tobacco etch 
virus (TEV) [1], tobacco vein mottling virus (TVMV) [4], pea seed-borne mosaic 
virus (PSbMV) [8], four strains of plum pox virus (PPV) [10, 12, 27], two strains 
of potato virus Y (PVY) [18, 28], papaya ringspot virus (PRSV) [31], pepper 
mottle virus (PeMV) [30], soybean mosaic virus (SMV) [7], potato virus 
A (PVA) [17], peanut stripe virus (PstV) [6] and Canadian strain of TuMV 
(TuMV-C) [15] have been published. For other strains of TuMV, some partial 
nucleotide sequences of the 3' terminal region including the CP and HC/Pro

* DDBJ/EMBL/GenBank accession number D83184.
protein genes have been reported [9, 13, 14]. Thus, the information regarding the extent of genomic variation among these TuMV isolates and strains is limited. In this paper, we describe the complete nucleotide sequence of Japanese strain of TuMV (TuMV-J) genomic RNA.

TuMV-J (strain 1) [22] was propagated on Brassica rapa L. cv. Hakatasuwari and purified by the method described by Choi et al. [3]. The genomic RNA of TuMV was extracted by the procedure described by Rosner et al. [20]. Two procedures were employed for the cDNA cloning of the TuMV-J gene. (1) First and second-strand cDNAs were synthesized from TuMV-J RNA using the reverse-transcription polymerase chain reaction method (RT-PCR) [16, 21]. The oligonucleotide primers were designed based on the sequence of TuMV-C [15] or TuMV-J determined in this study. (2) First-strand and second-strand cDNAs were synthesized from TuMV-J RNA according to the method of Gubler and Hoffman [9] using cDNA Synthesis System Plus (Amer- sham). The dsDNAs ligated into pBluescript II SK− were introduced into Escherichia coli XL1-Blue. The nucleotide sequence of the TuMV-J RNA was determined from five overlapping clones. The dsDNAs were sequenced by the dideoxynucleotide chain termination method [24] using modified T7 DNA polymerase (Sequenase, USB) or a Dye terminator sequencing kit (Amersham). Moreover, selected portions of TuMV-J RNA were sequenced using specific oligonucleotide primers and avian myeloblastosis virus reverse transcriptase. The nucleotide and deduced amino acid sequence analyses, similarity searches and alignments were carried out using DNASIS (Hitachi) and SDC-GENETYX computer programs (Software Development).

Among the large number of recombinant plasmids containing inserts of different sizes corresponding to various parts of the genomic RNA of TuMV-J, five overlapping recombinant plasmids were selected which represented the entire genome of TuMV-J RNA. Nucleotide sequences of these clones were determined at least twice in both orientations and no difference was observed in overlapping regions of these clones. The nucleotide sequences of the 5′ terminal region and a part of the cytoplasmic inclusion protein gene of TuMV-J were obtained by direct RNA sequencing. The nucleotide and deduced amino acid sequences of TuMV-J RNA will appear in DDBJ, EMBL and GenBank nucleotide sequence databases with the following accession number; D83184. The TuMV-J RNA genome was 9833 nucleotides in length, excluding a 3′ terminal poly(A) tail, and contains 129 and 209 nucleotide long of 5′ and 3′ non-coding regions, respectively. Computer analysis of the coding capacity of the TuMV-J RNA sequence revealed a single open reading frame (ORF) of 9 492 nucleotide starting at nucleotide position 130 and ending at a UGA termination codon at position 9622. Moreover, the sequence surrounding the AUG at position 130 fits with the plant consensus initiation sequence proposed by Lütcke et al. [11]. Therefore, the first AUG triplet at positions 130–132 is good candidate to be the initiation codon for translation. The deduced ORF would encode a polyprotein of 3164 amino acid residues with a calculated Mr of 3 57 700.