The 5′-terminal sequence of TMV RNA
Question on the polymorphism found in *vulgare* strain

Tetsuo Meshi, Masayuki Ishikawa, Nobuhiko Takamatsu, Takeshi Ohno and Yoshimi Okada

Department of Biophysics and Biochemistry, Faculty of Science, University of Tokyo, Hongo, Tokyo 113, Japan

Received 18 July 1983; revised version received 30 August 1983

The complete nucleotide sequence of TMV RNA (common strain) reported in [Proc. Natl. Acad. Sci. USA (1982) 79, 5818] its 5′-end to be represented by two variants which differed in length. We have tested that result and sequenced the 5′-terminal regions of two strains of TMV RNA (common strain OM and tomato strain L) using cloned cDNA copies. The results showed that the 5′-terminal region of the TMV genome is not polymorphic and that one of the two variants cited above represents a tomato strain but not the common strain.

Tobacco mosaic virus *Nucleotide sequence* *5′-Non-coding region* *Heterogeneity*

1. INTRODUCTION

Tobacco mosaic virus (TMV) has a single-stranded RNA genome of about 6400 nucleotides [1]. The complete sequence of the genome of a common strain (*vulgare*) of TMV was reported in [2]. They detected variability in the viral population especially in the 5′-terminal region [2]. It is widely accepted that there are some heterogeneities in the population of an RNA virus. However, the findings in [2] were not a base substitution or a heterogeneity in the length of poly(A) tract observed in [3–7], but the existence of two variants which could apparently be distinguished from each other [2].

Here, we analyzed the nucleotide sequences of the 5′-terminal regions of two closely related TMV strains (common strain OM and tomato strain L) [8]. Our results indicated that the 5′-terminal region of the TMV genome is not polymorphic and that one of the two variants reported in [2] represents a tomato strain but not the common strain.

2. MATERIALS AND METHODS

Synthesis of cDNA copies of TMV RNA and its cloning into *Escherichia coli* were described in [7,9]. For screening, Ω-RNA was prepared as in [10] and labeled at the 5′-end [11]. Colony hybridization was carried out as in [12].

The DNA sequence was determined as in [13]. Sequencing the 5′-labeled Ω-RNA was carried out as in [14,15]. The dideoxy chain terminating method was performed as follows. The genomic RNA and 5′-^{32}P-labeled restriction fragments were annealed in a sealed capillary in 0.3 M KCl, 50 mM Tris–HCl (pH 7.9 at 42°C) by heating at 90°C for 5 min and cooling slowly to 30°C. The annealing mixture was diluted and reverse transcribed in 60 mM Tris–HCl, 10 mM $MgCl_2$, 10 mM DTT, 60 mM KCl, 1 mM dNTPs and 100 Units/ml reverse transcriptase (Seikagaku Kogyo). The ratio of ddNTP (final 50 μM) to dNTP was 1:1 to 1:2. The reaction was started at 30°C and was then brought slowly to 42°C. After 30 min at 42°C, one dNTP was added to the mixture (final 1 mM) and incubation was continued for 15 min. After ethanol precipitation, the transcribed DNA was loaded onto the sequencing gel.

3. RESULTS

3.1. *Cloning and nucleotide sequence of the 5′-end of the genomic RNA*

Cloning of the cDNA copies of TMV RNA was

Published by Elsevier Science Publishers B.V.
00145793/83/$3.00

carried out as in [9]. The cDNA clones carrying the sequence of the 5'-extremity of the genomic RNA were selected by colony hybridization using the ^{32}P-labeled RNase T1-resistant Ω-RNA fragment derived from the 5'-end of the genomic RNA [10] as a probe. For sequence analysis we selected 4 clones of the OM (common) strain and 2 of the L (tomato) strain. Fig.1 shows the restriction map and the strategy developed for sequencing at the 5'-end of the genomic RNA.

The sequences of the cloned cDNA copies were determined as in [13]. The results are summerized in fig.2. No mismatch was found among the 4 OM clones or between the 2 L clones.

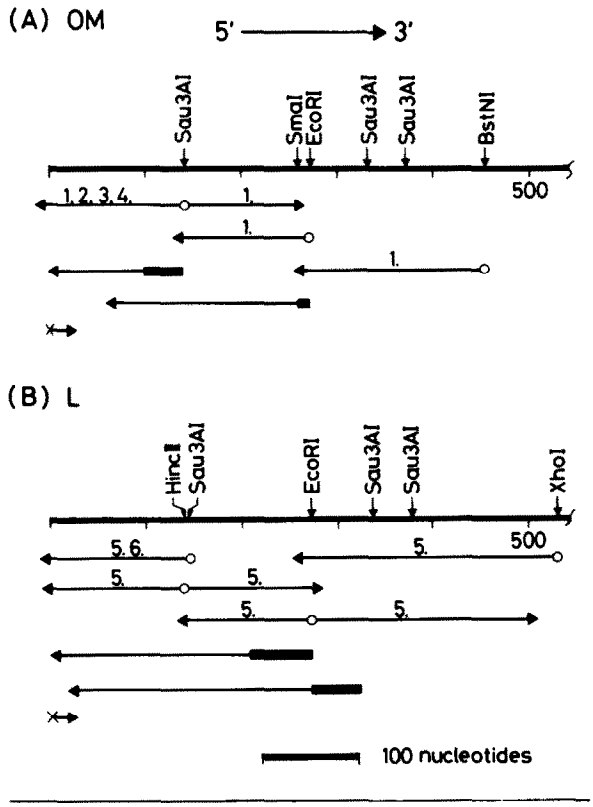

Fig.1. Restriction maps and strategies for sequencing of about 500 nucleotides from the 5'-end of TMV OM (A) and L (B) RNA. Terminal dots indicate the labeled ends of restriction fragments as used in [13]. Plasmids used are abbreviated by numbers above the thin arrows: (1) pOM-D100; (2) pOM-A4; (3) pOM-D98; (4) pOM-F92; (5) pL-D66; (6) pL-D82. Thick short bars at the ends of the thin arrows denote the restriction fragments used as primers for the chain terminating method. Terminal crosses indicate the labeled termini of Ω-RNAs.

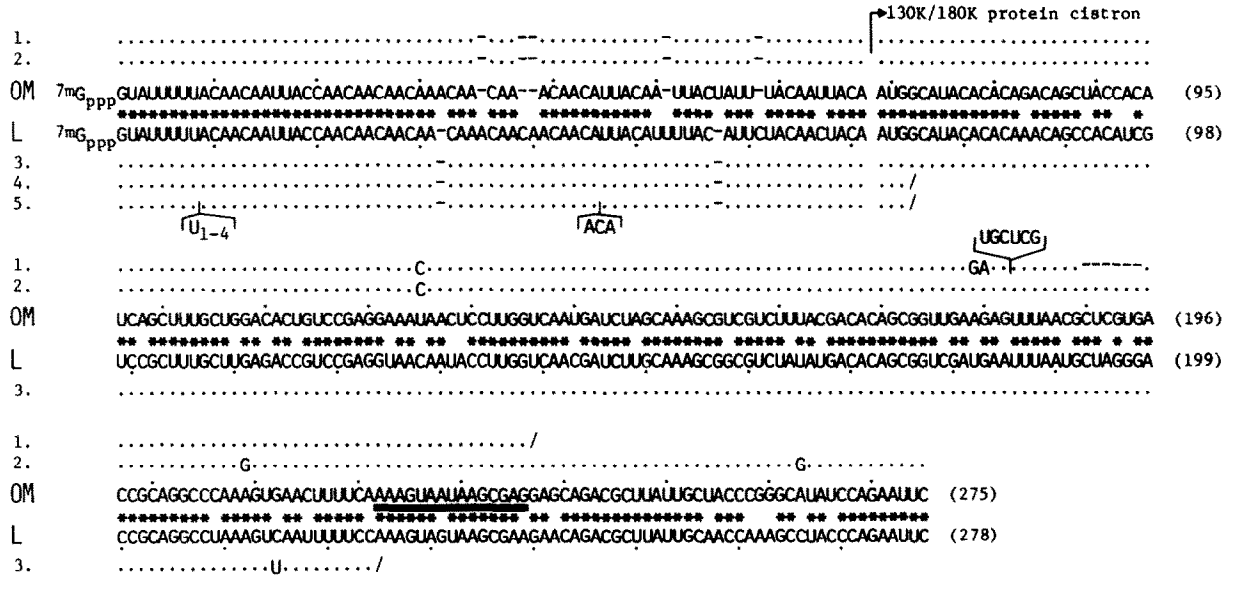

Fig.2. Comparison of the nucleotide sequences of the 5'-terminal regions of TMV RNAs. Common residues between OM and L strain are shown by an asterisk. Hyphens denote gaps inserted for alignment. Sequence 1 is the *vulgare* strain from [17,18]; (2,3) are, respectively, the shorter and longer variants of the *vulgare* strain from [2]; (4) is the SPS strain from [19]; (5) is the *dahlemense* strain from [21]. Sequences (1,2) are compared with that of the OM strain, and (3,4,5) with that of the L strain. Identical residues are indicated by a dot. The underlined region (residues 222–236 of the OM strain) indicates the complement of the primer used in [2] (see text).

3.2. *Further confirmation of the sequence*

To examine whether the nucleotide sequence determined using cloned cDNA copies reflected the genomic RNA sequence, we sequenced the genomic RNA directly by the dideoxy chain terminating method [16]. Short restriction fragments (fig.1) labeled by ^{32}P at one 5′-end were prepared as primers, annealed with the genomic RNA and reverse transcribed. Fig.3 shows an example of the sequencing ladder. If there was any heterogeneity, bands with the same electrophoretic mobility would appear in at least two tracks. We did not find any confusing pattern, except some ambiguities caused by premature termination of reverse transcriptase (fig.3). The results perfectly match those determined using cloned cDNAs. The sequence shown in fig.2 therefore represents the major population of the genomic RNA. This indicates clearly that the 5′-terminus of the TMV RNA is well-conserved and that there is little heterogeneity, if any.

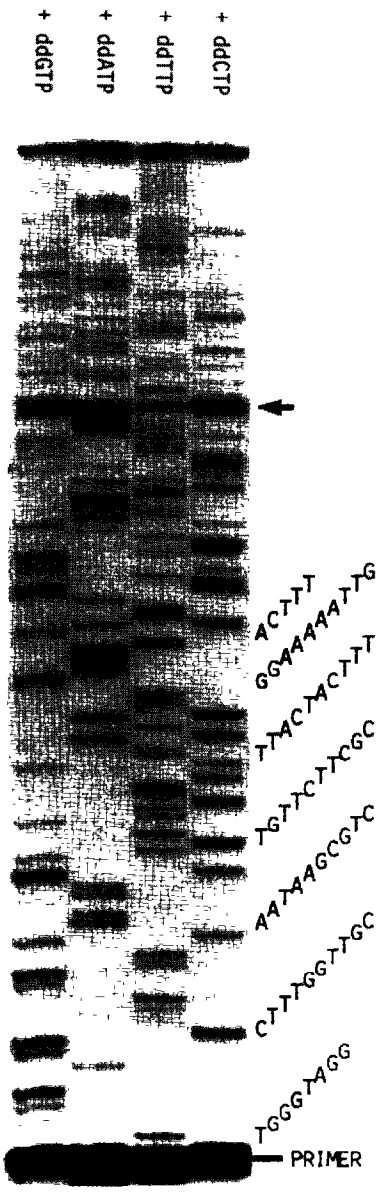

Fig.3. Sequencing ladder of chain terminating method. The primer used is the *Sau*3AI–*Eco*RI fragment of pL-D66 (see text). The short arrow shows a site of the premature termination of the reverse transcriptase, where a band appears even in the absence of any dideoxynucleotide.

4. DISCUSSION

Two TMV strains used in this work, the common strain OM and the tomato strain L, were isolated in Japan and are thought to correspond to the *vulgare* (or U1) and *dahlemense* strains, respectively. Comparing about 1600 nucleotides at the 3′-end, the homology between OM and *vulgare* is about 98% and that between OM and L is about 75% [7]. In the 5′-terminal 275 residues presented in fig.2, the homology between OM and L is about 84%, slightly higher than observed in the 3′ noncoding region.

It has been reported [2] that the 5′-end of TMV (residue 1–221) was polymorphic and that there were two variants (a short and longer one). Comparing our data with those in ([2], fig.2), we concluded that the longer variant was not the common strain but reflected the sequence of a tomato strain, which was probably contaminated during inoculation or preparation, for the following reasons. The homology between the 5′-end region of the L strain and the longer variant is extensively high. Only one mismatch was found in total of 221 residues (fig.2), which indicates that the two are the same strain. We could observe no heterogeneities in our TMV preparations, even using, as primer, the fragment complementary to residues 260–274 of OM RNA or to residues 278–337 of L RNA derived from the region where the polymorphism was not detected in [2]. This means that such a molecule as a hybrid between a tomato and a common strain does not exist. When cloning the 5′-end of the genomic

RNA, authors in [2] used a synthetic nucleotide as primer. The primer was, however, a complement not only of the common strain RNA (residue 222–236) but also of the tomato strain RNA (residue 225–238, fig.2). This would be the reason why polymorphism at the 5′-end found in [2] was restricted only to residues 1–221. As regards the downstream region (residues 222–455), the homology between OM and *vulgare* is 98% (not shown). On the other hand, the homology in that region between *vulgare* and L is about 85% (not shown). Therefore, the genomic sequence in [2] must represent the sequence of the common strain except for the 5′ polymorphism.

The sequence data reported in [17,18], are in agreement with the OM sequence, except for changes in the region of residues 179–195 (fig.2). However, these changes might be explained by mis-ordering of the RNase T1 oligonucleotides.

The 5′ non-coding region of the SPS strain has been published [19] and it was in agreement with the sequence of the L strain (fig.2). This indicates that the SPS strain is closely related to the tomato strain.

Is there true heterogeneity in the non-coding region? As for the 3′ non-coding region, the same result has been obtained in 3 laboratories in spite of the use of independent isolates and stocks of the common strain [2,7,20]. Our results also suggest that there is little, if any, heterogeneity in the 5′ non-coding region. Although replication of TMV still remains to be well-characterized, it seems reasonable to assume that conservation of the 3′ and 5′ non-coding regions must be important for this step of virus propagation.

ACKNOWLEDGEMENT

This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science and Culture, Japan.

REFERENCES

[1] Hirth, L. and Richards, K.E. (1981) Adv. Virus Res. 26, 145–199.
[2] Goelet, P., Lomonossoff, G.P., Butler, P.J.G., Akam, M.E., Gait, M.J. and Karn, J. (1982) Proc. Natl. Acad. Sci. USA 79, 5818–5822.
[3] Domingo, E., Sabo, D., Taniguchi, and Weissmann, C. (1978) Cell 13, 735–744.
[4] Ysebaert, M., Van Emmelo, J. and Fiers, W. (1980) J. Mol. Biol. 143, 273–287.
[5] Ahlquist, P., Luckow, V. and Kaesberg, P. (1981) J. Mol. Biol. 153, 23–38.
[6] Meshi, T., Ohno, T. and Okada, Y. (1982) Nucleic Acids Res. 10, 6111–6117.
[7] Takamatsu, N., Ohno, T., Meshi, T. and Okada, Y. (1983) Nucleic Acids Res. 11 3767–3778.
[8] Gibbs, A.J. (1977) CMI/AAB Descrip. Plant Viruses no. 184.
[9] Okayama, H. and Berg, P. (1982) Mol. Cell. Biol. 2, 161–170.
[10] Mandeles, S. (1968) J. Biol. Chem. 243, 3671–3674.
[11] Meshi, T., Takamatsu, N., Ohno, T. and Okada, Y. (1982) Virology 118, 64–75.
[12] Grunstein, M. and Hogness, D.S. (1975) Proc. Natl. Acad. Sci. USA 72, 3961–3965.
[13] Maxam, A.M. and Gilbert, W. (1980) in: Methods in Enzymology (Grossman, L. ed) vol. 65, pp. 499–560, Academic Press, New York.
[14] De Wachter, R. and Fiers, W. (1972) Anal. Biochem. 49, 184–197.
[15] Donis-Keller, R., Maxam, A.M. and Gilbert, W. (1977) Nucleic Acids Res. 4, 2527–2538.
[16] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463–5467.
[17] Jonard, G., Richards, K., Mohier, E. and Gerlinger, P. (1978) Eur. J. Biochem. 84, 521–531.
[18] Richards, K., Guilley, H., Jonard, G. and Hirth, L. (1978) Eur. J. Biochem. 84, 513–519.
[19] Konarska, M., Filipowicz, W., Domdey, H. and Cross, H.J. (1981) Eur. J. Biochem. 114, 221–227.
[20] Guilley, H., Jonard, G., Kukla, B. and Richards, K.E. (1979) Nucleic Acids Res. 6, 1287–1308.
[21] Kukla, B.A., Guilley, H.A., Jonard, G.X., Richards, K.E. and Mundry, K.W. (1979) Eur. J. Biochem. 98, 61–66.