Complete nucleotide sequence of the genome of an apple isolate of apple chlorotic leaf spot virus

K. Sato, N. Yoshikawa* and T. Takahashi

Plant Pathology Laboratory, Faculty of Agriculture, Iwate University, Ueda 3-chome, Morioka 020, Japan

The complete nucleotide sequence of the genome of an apple isolate of apple chlorotic leaf spot virus (ACLSV-A) was determined. The genome is 7552 nucleotides excluding the poly(A) tail and contains three open reading frames (ORFs 1, 2 and 3), encoding proteins with Mr values of 216503 (216.5K), 50453 (50.4K) and 21394 (21.4K), respectively. Nucleotide sequence comparisons between ACLSV-A and the previously sequenced ACLSV from plum (ACLSV-P) showed that the sequence identity at the nucleotide level was 79.8%. Amino acid sequence identities of ORFs 1 and 2 between both isolates were 88.4% and 79.9%, respectively. The 21.4K protein encoded by ORF 3 of ACLSV-A had an amino acid sequence identity of 88.6% with the 28.3K protein encoded by ORF 3 of ACLSV-P. Immunoblot analysis of the 21-4K protein expressed in Escherichia coli showed that this protein is the coat protein of ACLSV-A.

Apple chlorotic leaf spot virus (ACLSV) is distributed worldwide in fruit trees including apple, peach, pear, plum, cherry and apricot (Lister, 1970; Németh, 1986). In Japan, this virus is one of the causative agents of apple topworking disease and induces lethal decline in apple trees grown on Maruba kaido (Malus prunifolia ringo) rootstocks (Yanase, 1974). ACLSV has very flexuous filamentous particles, approximately 600 to 700 nm in length and contains a polyadenylated plus-sense, ssRNA with an Mr of 2-48 x 10^6 and a single coat protein of 22K (Yoshikawa & Takahashi, 1988). At present, this virus is classified into the closterovirus group based on its particle structure (Francki et al., 1991). Recently, German et al. (1990) reported the nucleotide sequence of the genome of ACLSV isolated from plum (Prunus domestica) (ACLSV-P) in France.

In this paper, we report the complete nucleotide sequence of the genome of ACLSV isolated from apple (ACLSV-A) in Japan and the comparisons of the nucleotide sequence and the deduced amino acid sequence with those of ACLSV-P.

ACLSV-A (isolate P-205), originally isolated from an apple tree (Yanase, 1974), was propagated in Chenopodium quinoa and purified as described previously (Yoshikawa & Takahashi, 1988). Viral RNA was extracted from purified virus by dissociation in 2% SDS for 15 min, followed by extraction with phenol-chloroform, precipitated from 70% ethanol and suspended in water. First and second strand cDNAs were prepared from 2μg of ACLSV-A RNA according to Gubler & Hoffmann (1983) using oligo(dT) or random hexanucleotides as primers. The dsDNAs were blunt-ended using T4 DNA polymerase and ligated into the EcoRV site of Bluescript KS+ (BS) or dC-tailed and annealed to dG-tailed, PstI-cut pUC9 (Yoshikawa et al., 1988). The resulting plasmids were used to transform competent Escherichia coli DH5α cells.

The cDNA clones or subclones produced after the ligation of restriction fragments into BS were digested with exonuclease III and mung bean nuclease (Henikoff, 1984) and the resulting deletion mutants were used for sequencing by the dideoxynucleotide chain termination method (Sanger et al., 1977) using Sequenase version 2.0 (United States Biochemicals). Each base was determined by sequencing at least two independent clones. More than 80% of the genome was sequenced in both directions and dITP was used to resolve ambiguities resulting from compression.

The sequence of the 5'-terminal region of ACLSV-A RNA was determined by extending a synthetic primer (5' AGGGCGTTACCTCAATCTG 3') complementary to nucleotides (nt) 38 to 55 of the viral RNA using reverse transcriptase and terminal deoxynucleotidyl transferase (Deborde et al., 1986). All nucleotide sequence data were collected and analysed using the program GENETYX version 8.0 (Software Development Co. Ltd).

The genome of ACLSV-A consists of 7552 nt ex-
Fig. 1. The complete nucleotide sequence of the ACLSV-A genome and the amino acid sequences of three ORFs. The amino acid sequences of the conserved motifs associated with the helicase and the RNA polymerase are underlined.

including the 3' poly(A) tail (Fig. 1), three nt fewer than that of ACLSV-P reported by German et al. (1990) (Table 1). Analysis of the putative open reading frames (ORFs) of the nucleotide sequence in both positive and negative strands showed that the genome of ACLSV-A contains three ORFs in the positive strand (Fig. 1).
Table 1. Comparisons between ACLSV-A and ACLSV-P genomes

<table>
<thead>
<tr>
<th></th>
<th>ACLSV-A</th>
<th>ACLSV-P</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length excluding poly(A) tail</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole genome</td>
<td>7552 nt</td>
<td>7555 nt</td>
<td>79.8%</td>
</tr>
<tr>
<td>5' non-coding region</td>
<td>151 nt</td>
<td>151 nt</td>
<td>92.7%</td>
</tr>
<tr>
<td>3' non-coding region</td>
<td>187 nt</td>
<td>190 nt</td>
<td>81.2%</td>
</tr>
<tr>
<td>M_r of ORF products</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORF 1</td>
<td>216503</td>
<td>216559</td>
<td>88.4%</td>
</tr>
<tr>
<td>ORF 2</td>
<td>50453</td>
<td>50831</td>
<td>79.9%</td>
</tr>
<tr>
<td>ORF 3</td>
<td>21394</td>
<td>28376</td>
<td>88.6%</td>
</tr>
</tbody>
</table>

ACLSV-A (S715) A CGG CAT A GC C AT C AT GATAA

ACLSV-P (S712) A CGG CAT G C G A C AT GATAA

Fig. 2. The nucleotide sequences surrounding the ORF 2 initiation codons of ACLSV-A and ACLSV-P. Identical nucleotides are shown by asterisks. Second ATG codons within ORF 2 are underlined.

ORF 1, preceded by a 151 base leader sequence, starts at AUG (positions 152 to 154) and ends at UGA (positions 5807 to 5809) and encodes a protein with an M_r of 216503 (216.5K). The ORF 1 protein contains two motifs associated with the helicase (GxxxGxGKS) and the RNA polymerase (GxxxTxxxNT, GDD) at positions 1060 to 1067 and 1694 to 1728, respectively (Fig. 1). The overall similarity of the ORF 1 protein between ACLSV-A and ACLSV-P is 88.4% (Table 1). However, the regions surrounding the above two motifs were highly conserved between both isolates and their identities were 97.6% (positions 1058 to 1305 for ACLSV-A and 1057 to 1304 for ACLSV-P) and 97.7% (positions 1637 to 1767 for ACLSV-A and 1636 to 1766 for ACLSV-P), respectively.

ORF 2 begins at AUG (positions 5727 to 5729) and terminates at UGA (positions 7098 to 7100) to yield a protein with an M_r of 50453 (50.4K) (Fig. 1). The position of the initiation codon of the 50.4K protein is six nt downstream from the corresponding position in ACLSV-P (Fig. 2). This location of the initiation codon was the same in the three different cDNA clones sequenced. The 50.4K protein of ACLSV-A, as did that of ACLSV-P, showed an amino acid similarity of 22.4% with the gene I protein of cauliflower mosaic virus (Franck et al., 1980), and the identity of the ORF 2 protein between ACLSV-A and ACLSV-P was 79.9% (Table 1).

ORF 3 contains an AUG (positions 6784 to 6786) and ends at UGA (positions 7363 to 7365) encoding a protein with an M_r of 21394 (21.4K) which is dissimilar to the 28.3K protein encoded by ORF 3 of ACLSV-P. This discrepancy is explained as follows. The initiation codon (AUG, positions 6616 to 6618) is located in the ACLSV-A genome at the almost same position as the AUG codon of the 28.3K protein of ACLSV-P, but the termination codon (UGA, positions 6670 to 6672) exists between the first AUG and the second AUG (positions 6784 to 6786) which is the initiation codon of the 21.4K protein (Fig. 3). The identity of ORF 3 proteins between both isolates was 88.6% (Table 1).

In in vitro translation experiments reported previously...
ACLSV-A agrees with that of the ACLSV-A coat protein and several minor proteins of which the 23K protein (Yoshikawa & Takahashi, 1989), ACLSV-A RNA responding to nt positions 6781 to 6798 and containing cultures grown in the presence of IPTG and analysed by immunoblotting using ACLSV-A antiserum as described previously (Yoshikawa et al., 1992). This protein is slightly larger than coat protein from purified virus and demonstrates that the 21-4K protein encoded by ORF 3 is the coat protein of ACLSV-A, supporting the results obtained before following the determination of the partial amino acid sequence of the ACLSV-P coat protein (German et al., 1990).

At present, ACLSV is classified into the closterovirus group, based mainly on its particle structure (Francki et al., 1991). This group is composed of heterogeneous virus species with different particle lengths, transmission modes and tissue localization of the virus particles in infected material. The genome organization of ACLSV is quite different from that of the 3'-terminal genomic region of beet yellows virus (Agranovsky et al., 1991), the type member of the closterovirus group, indicating that ACLSV should be classified into a new virus group. Further investigations on the genome organization of closteroviruses will provide valuable information on the classification of this group.

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

References


(Received 8 March 1993; Accepted 29 April 1993)