SHORT COMMUNICATION

M. J. GIBBS,*† J. I. COOPER,* and P. M. WATERHOUSE†

The genomic sequence of an Australian isolate of carrot mottle umbravirus (CMoV-A) was determined from cDNA generated from dsRNA. This provides the first data on the genome organization and phylogeny of an umbravirus. The 4201-nucleotide genome contains four major open reading frames (ORFs). Analysis suggests that ORF2 encodes an RNA-dependent RNA polymerase, that ORF4 encodes a movement protein, and that the virus has no coat protein gene. The functions of ORFs 1 and 3 remain unknown. ORF2 is probably translated following ribosomal frameshifting. ORFs 3 and 4 are probably translated from a subgenomic mRNA. Sequence comparisons showed CMoV-A to be closely related to pea enation mosaic RNA2 (PEMV-RNA2), but also to have affinities with the Bromoviridae. These findings shed light on the relationships between the luteoviruses, PEMV, and the umbraviruses and on the relationships between the carmo-like viruses and the Bromoviridae.

In nature umbraviruses have been found only in plants co-infected with a luteovirus from which they are distinguished because they are mechanically transmissible, whereas luteoviruses are transmitted only by aphids (1, 2). The Umbravirus genus was defined in 1995 (2) primarily on the basis of two features. First, umbraviruses depend on an interaction with the co-infecting luteoviruses for their transmission by aphids (3, 4). Second, although umbraviruses have monopartite single-stranded RNA genomes (3, 5), infected plants contain two abundant double-stranded RNA species of about 4.5 and 1.4 kilo base-pairs (kbp) (2).

Carrot mottle umbravirus (CMoV), the type species of the Umbravirus genus, is found in association with carrot red leaf luteovirus (CRLV) (6, 7). An Australian isolate of CMoV (CMoV-A) was obtained from carrots growing in a commercial plot in New South Wales. The virus was identified by its host range, by its dependence on CRLV for aphid transmission, and by the symptoms and double-stranded RNA (dsRNA) profile (Fig. 1, top) of infected plants.

As yet, no nucleotide sequence has been reported for an umbravirus. With a view to understanding the genome organization of umbraviruses and their relationships with other plant viruses we determined the genomic sequence of CMoV-A. The virions produced by transcapsidation were believed to be a poor source of genomic RNA (8). Hence, CMoV-A dsRNA, purified from total nucleic acids (9) from infected Nicotiana clevelandii, was used for cDNA synthesis (10–12). Initially, random hexamer primers were used to prime cDNA synthesis and the products were cloned by a shotgun method (13). Later, sequence-specific primers were used and cDNA was cloned after PCR amplification (14). cDNA to the genomic 5′ terminus was obtained by ligating a 3′-NH₂ blocked primer to the dsRNA and using a second primer, complementary to the blocked primer, in cDNA synthesis and PCRs (12). cDNA to the genomic 3′ terminus was obtained by polyadenylating the dsRNA and using an oligo(dt) primer in cDNA synthesis and PCRs. The compilation of sequences from 100 CMoV-A cDNA clones gave a single contiguous sequence of 4201 nucleotides (GenBank Accession No. U57305). This sequence included the terminal nucleotides as estimated by primer extension analysis using primers A-5′ (CCGAGCTGGTGTTGACC) and B-3′ (GTTGGAGCCAAAACCTCGGGTG) (12). This is the first example of the cloning and sequencing of a single-stranded RNA viral genome entirely from dsRNA.

All 4201 nucleotide residues were determined from cDNA sequenced in both orientations. Except for the 5′ terminal 126 nucleotides and the 3′ terminal 101 nucleotides, the sequence was determined from two or more independently generated clones. As listed in the GenBank entry, at 16 positions the identity of a nucleotide differed when read from different clones spanning the
are also strong stem-loop structures 5' and 3' of the likely CMoV-A frameshifting slippery site (Fig. 3A) similar to those required for frameshifting by luteoviruses and PEMV-RNA2 (21, 24), but no pseudoknot was detected. The translation of CMoV-A ORF1 and ribosomal frameshifting into ORF2 at the likely frameshifting site would yield a 98-kDa protein.

The AUG of ORF3 (nucleotide 2791) has a relatively poor translational context (GUCAUGAA) (25) and lies 28 nucleotides 5' of the AUG of ORF4 (nucleotide 2819) which has a more optimal context (GCGAUGGC). This, and the genomic location of these ORFs, suggests they are probably translated from a single subgenomic mRNA (26). The small (1.4 kbp) dsRNA found in CMoV-A-infected plants is the correct size to be this subgenomic mRNA.

CMoV-A RNA species were transferred from agarose gels to nylon membranes and incubated with radiolabeled probes (27) prepared from CMoV-A cloned cDNAs. These cDNAs represented three genomic regions: A, 5' terminal nucleotides 1 to 819; B, nucleotides 2852 to 3301, which encodes part of ORFs 3 and 4; and C, 3' terminal nucleotides 3670 to 4201. The probe to region A hybridized only with the large (4.5 kbp) CMoV-A dsRNA, whereas the probes to regions B and C hybridized to both the large and the small CMoV-A dsRNAs (Fig. 1, bottom). This suggests that the small CMoV-A dsRNA species represents a dsRNA form (28) of a 3' coterminal subgenomic mRNA.

Database searches (29) detected similarity between CMoV-A ORF2 and the RdRp genes from viruses from the carmovirus, dianthovirus, luteovirus, machlomovirus, necrovirus, and tombusvirus genera (the carmo-like grouping) (30). This similarity was shown to represent homology when standard deviation scores (SD scores) (31, 32) ranging from 25 to 42 were obtained for alignments of the amino acid sequences encoded by these genes (GenBank and EMBL Accession Nos.: X07653,

![Figure 1](image.png)  
**FIG. 1.** (Top) The dsRNA profile of CMoV-A-infected Nicotiana clevelandii (lane 2). Lane 1 contained bacteriophage λ DNA treated with the restriction endonucleases EcoRI and HindIII. The lengths of the dsRNA species were estimated to be 4.5 and 1.4 kbp (A and B). (Bottom) A Northern blot of dsRNA species from CMoV-infected N. clevelandii probed with radiolabeled DNA of clone 684 (nucleotides 3670 to 4201). Lane 1 was loaded with the large dsRNA species (top panel, species A) purified from a single band on an agarose gel. Lane 2 was loaded with the small dsRNA species (top panel, species B) also purified from a single band on an agarose gel. Both dsRNA species migrated as two bands (A1 and A2, and B1 and B2), presumably by separating into dsRNA and ssRNA forms, under denaturing conditions.

![Figure 2](image.png)  
**FIG. 2.** The genome plan of CMoV-A. Major ORFs are shown as boxes and arranged according to the reading frame used with the first reading frame at the top. The predicted molecular weights of the proteins encoded by these ORFs are shown in parentheses. The molecular weight shown for ORF2 is that calculated for ORFs 1 and 2 assuming −1 ribosomal frameshifting between the ORFs as described in the text. The horizontal line beneath the ORFs represents the full length of the genome.
both RNAs spread systemically (33, 38). However, when separated, only PEMV-RNA2 can establish a systemic infection (33, 38). Nevertheless, PEMV-RNA1 can infect protoplasts (38), suggesting that PEMV-RNA2 complements PEMV-RNA1 by allowing systemic spread (33). Considering our analysis, we suggest that the protein translated from PEMV-RNA2 ORF4 has this complementary role.

Database searches failed to detect any similarity between the amino acid sequences encoded by CMoV-A and any viral coat protein sequence, suggesting that CMoV-A, like PEMV-RNA2 (33), probably does not produce its own virions. However, on the basis of electron microscopy it has been suggested that umbraviruses have pleiomorphic enveloped virions (39). It is interesting to notice that these structures were identified budding from the tonoplast into the vacuole. Strikingly similar vesicles, thought to be sites of replication, are found budding from the tonoplast into the vacuole in CMV-infected plants (40, 41). Considering the homology between CMoV-A ORF4 and the movement protein gene of CMV, we suggest that the structures found in umbravirus-infected plants are not particles but may be replication sites in which a common movement protein plays some role.

Given the affinities between the ORFs in the genome of CMoV-A and PEMV-RNA2, a search was made for similar nucleotide sequences and RNA structures in these two molecules. In addition to similarities in sequence and RNA secondary structure at the likely frameshifting sites (Fig. 3A) (24), the last 6 nucleotides at the genomic 3′ termini are identical, and there is a 15-nucleotide stretch of identity (CGGGCAUAUAAUAGG) starting about 65 nucleotides 5′ of the genomic 3′ termini. Similar stem–loop structures were also predicted to form close to the genomic 3′ termini (Figs. 3B and 3C). These stem–loop structures closely match stem–loop structures predicted to form at the 3′ termini of both genomic segments of the dianthoviruses RCNMV and SCNMV (19, 42). It is likely that some or all of these features identified close

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A Comparison between the Proteins Encoded by the Four ORFs of CMoV-A and Their Homologues Encoded by PEMV-RNA2</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ORF</th>
<th>CMoV-A (kDa)</th>
<th>PEMV2 (kDa)</th>
<th>Amino acid percentage identity</th>
<th>Similarity estimate (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>37</td>
<td>33</td>
<td>30</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>64</td>
<td>65</td>
<td>57</td>
<td>56</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>25</td>
<td>21</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>27</td>
<td>39</td>
<td>27</td>
</tr>
</tbody>
</table>

Note. Predicted molecular weights are shown together with the percentage identity and a similarity estimate (31, 32) for each homologous pair. The entire coding sequence of ORF2 from each virus, disregarding the initiation codon (AUG), was used in these calculations.
to the genomic 3' termini are important for genomic RNA replication.

PEMV is currently recognized as a bipartite virus and the only species of the genus Enamovirus (43). However, current data show that PEMV-RNA1 is very similar to some luteoviruses (38, 44), and our comparisons show that PEMV-RNA2 is very similar to an umbravirus. These affinities, together with data showing both PEMV-RNA1 and PEMV-RNA2 to be capable of independent replication (33, 38) support the proposal (33) that PEMV is composed of two viruses. If PEMV-RNA2 is an umbravirus, then the features we have identified from sequence comparisons are likely to be common to other umbraviruses.

ACKNOWLEDGMENTS

We thank P. K. Keese of the CSIRO Division of Plant Industry for useful discussions, and J. N. Clarke of the University of Southampton Medical School for providing a 3'-blocked primer. During the course of this work M. J. Gibbs was a student of the University of Oxford Department of Plant Sciences and the Australian National University Department of Botany and was the recipient of a Committee of Vice Chancellors Overseas Research Students Award.

REFERENCES