Conservation of the 3′ terminal nucleotide sequence in five carlaviruses

M.Teresa M.Haylor, Alan A.Brunt[1] and Robert H.A.Coutts*
Department of Biology, Imperial College of Science, Technology, and Medicine, Prince Consort Road, London SW7 2BB and [1]Institute of Horticultural Research, Littlehampton, West Sussex BN17 6LP, UK

Submitted September 19, 1990 EMBL accession no. X55897

Carnation latent virus (CLV) is the type member of the carlavirus group (1), a group with positive sense RNA genomes of M_r $2.3-3.0\times10^6$ which are polyadenylated at the 3′ terminus. The nucleotide sequence of the 3′ proximal regions of four carlaviruses: potato virus M (PVM), potato virus S (PVS), lily symptomless virus (LSV) and *Helenium* virus S (HelVS) have recently been presented (2, 3, 4, 5). Using a similar procedure for cloning, employing oligo (dT) priming of first strand cDNA synthesis, we have generated a library of 3′ terminal clones of CLV RNA which have been sequenced by the dideoxynucleotide termination procedure. An examination of the 3′ sequence of PVM (2), PVS (3), LSV (4) and HelVS (5) RNAs and a comparison with the 3′ nucleotide sequence of CLV RNA reveals some striking similarities. These similarities are conserved in all five viruses in the central region of a putative 3′ terminal open-reading frame (ORF) coding for polypeptides of 12.6, 10.8, 11.6, 10.7 and 16 KDa for respectively HelVS, PVM, CLV, PVS and LSV (Figure 1). Within this region of high conservation four Cys residues which conform to a consensus sequence for a putative 'zinc-finger' nucleic acid-binding domain (6) are also conserved. All five proteins contain a high proportion of positively charged residues with CLV containing 15 Arg and 9 Lys residues and 10 residues of Glu and Asp. This 3′ terminal ORF appears to be unique to carlaviruses and sets them apart from other virus groups but no direct evidence for its expression *in vivo* or *in vitro* has been presented (2, 3, 4, 5). Analysis of the 3′ terminal non-coding regions of the five RNAs using the UWGCG GAP program (7) revealed over 60% homology between the nucleotide sequences of CLV compared with the other four viruses (Figure 2). This 3′ region of the viral RNAs has the capacity to be folded into a series of stem-loop structures. Such strict sequence conservation at the 3′ terminus of carlaviruses may have implications with regard to RNA secondary structure and configuration and might be important in RNA replication. However the putative polyadenylation signal AATAAA present in the terminal sequence of HelVS (5) and PVM RNAs (2) is not conserved in all of the viruses.

ACKNOWLEDGEMENT

Teresa Haylor is in receipt of an SERC-CASE studentship.

REFERENCES

1. Koenig,R. (1982) CMI/AAB Descriptions of Plant Viruses No. 259.
2. Rupasov,V.V. *et al.* (1989) *J. Gen. Virol.* **70** 1861–1869.
3. Mackenzie,D.J. *et al.* (1989) *J. Gen. Virol.* **70**, 1053–1063.
4. Memelink,J. *et al.* (1990) *J. Gen. Virol.* **71**, 917–924.
5. Foster,G.D. *et al.* (1990) *J. Gen. Virol.* **71**, 1877–1880.
6. Klug,A. and Rhodes,D. (1987) *Trends In Biochemical Sciences* **12**, 465–469.
7. Devereaux *et al.* (1984) *Nucl. Acids Res.* **12**, 387–395.
8. Schwartz,R.M. and Dayhoff,M.O. (1978) in Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*. Nat. Biomed. Res. Found., Washington **5(Supp. 3)**, 353–358.

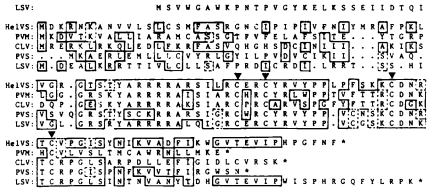

Figure 1. Alignment of the predicted amino acid sequences encoded by HelVS ORF 12.6 KDa, PVM ORF 10.8 KDa, CLV ORF 11.6 KDa, PVS ORF 10.7 KDa and LSV ORF 16 KDa using the GAP program (7). Identical amino acid residues between the ORFs according to Schwartz and Dayhoff (8), are boxed. The putative 'zinc-finger' domains are indicated by black triangles. Termination codons are shown as asterisks.

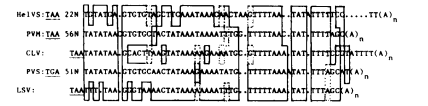

Figure 2. Sequence homologies in the 3′ non-coding regions of HelVS, PVM, CLV, PVS and LSV RNAs (shown as DNA). Sequences were aligned using the GAP program (7). Common nucleotides to the viruses are boxed. The termination codon of the terminal ORF is shown underlined for all five viruses. In the case of HelVS, PVM and PVS the number of nucleotides (N) between the termination codons and the start of the analysed sequence is indicated. All five RNAs terminate in a polyadenylated tail.

* To whom correspondence should be addressed