A 1.5 kb sequence homology in 3'-terminal regions of RNA-1 and RNA-2 of a birch isolate of cherry leaf roll nepovirus is also present, in part, in a rhubarb isolate

N. W. Scott,1* J. I. Cooper,1 Y. Y. Liu1 and C. U. T. Hellen2

1NERC Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR, U.K. and 2Department of Microbiology, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794-8621, U.S.A.

A series of cDNA clones has been made from the birch isolate of cherry leaf roll nepovirus. Restriction enzyme analysis and sequencing showed that at the 3' end, RNA-1 and RNA-2 are identical for 1.5 kb. Also a 0.7 kb 3' end homology exists between the birch and rhubarb isolate. These sequences do not seem to code for any proteins; however, the sequence conservation points to a role in virus replication.

Cherry leaf roll nepovirus (CLRV) has a bipartite genome composed of positive-sense ssRNA. The RNA-1 and RNA-2 are of Mf 2.82 x 10^6 and 2.29 x 10^6 respectively (Murant et al., 1981). They have a genome-linked protein (VPg) at the 5' end and are polyadenylated at the 3' end (Jones, 1985; Hellen & Cooper, 1987). Isolates of CLRV from different plant hosts are all of a similar size (protein, capsid and RNA); however, they show some structural diversity in that they can be distinguished antigenically (Jones, 1976) and by cDNA-RNA hybridization (Massalski & Cooper, 1986). Massalski (1984), in his hybridization studies, also found evidence consistent with the possibility that RNA-1 and RNA-2 of a CLRV isolate from birch (I2) have up to 2 kb of sequence in common. Such duplications have been reported in other viruses, for instance in the tobraviruses pepper ringspot (459 nucleotides) and tobacco rattle (1099 nucleotides) (Bergh et al., 1985; Angenent et al., 1986) and, in particular, in the nepovirus tomato ringspot (TomRSV). TomRSV like CLRV has a large RNA-2, so it is particularly interesting that Rott et al. (1988), after restriction analysis and sequencing of cloned DNA representing parts of the TomRSV genome, showed there was a substantial sequence duplication (approx. 1.5 kb) in the 3' ends of RNA-1 and RNA-2 of this virus.

Here we present sequence data for the 3' ends of two of the most distinct CLRV isolates, I2 and rhubarb (R) and show that a 3' end duplication occurs in CLRV I2. The virus genome was cloned, utilizing cDNA made from isolated RNA. The general methods of virus propagation and RNA extraction used have been described more fully elsewhere (Jones & Mayo, 1972; Massalski & Cooper, 1986). I2 and R RNA was isolated using the Pharmacia mRNA isolation system; in both cases RNA-1 and RNA-2 were polyadenylated. Initially mixed I2 RNA-1/RNA-2 populations were used as templates for oligo(dT)12,18-primed cDNA synthesis (Pharmacia). However, irrespective of RNA pretreatment (e.g. heat denaturation prior to cDNA synthesis) the clones obtained were short, no longer than 1.5 kb. All the clones obtained had similar restriction enzyme maps and all hybridized to both RNA-1 and RNA-2 components of I2 in Northern blot analysis. These data suggested that the 3' ends of RNA-1 and RNA-2 were homologous. To confirm this it was necessary to obtain clones that overlapped this common sequence and the RNA-1 and RNA-2 unique regions. A random priming strategy for cDNA synthesis was used (Boehringer Mannheim cDNA Synthesis Kit) and the cDNA was ligated using EcoRI/NotI adaptors into pT7T3 18U (Pharmacia). The plasmid was transferred into Escherichia coli (strain XL1-blue) using a method developed by Merrick et al. (1987), except the cells suspended in 50 mM-CaCl2 were kept on ice for a 10 min period prior to and after the addition of plasmid DNA. The cells were frozen in liquid nitrogen for 2 min, thawed at 37 °C, incubated for 2 h in LB medium without selection and then plated onto LB agar containing 50 µg/ml ampicillin. Colonies containing plasmid with inserts were identified using the X-Gal selection system, the inserts were sized and characterized by restriction enzyme analysis. Eleven of the largest random primed clones and one oligo(dT)12,18-primed clone (3X) (Fig. 1) were used to make photobiotin-labelled probes (McInnes et al., 1987). These were used in a Northern blot analysis against I2 RNA (Fig. 1). The conditions used were as described by Anderson & Young (1985) and Wahl et al. (1987) except for the stringency wash where filters were incubated in 0.15 x SSC, 0.1%
Short communication

Fig. 1. Northern blot analysis of RNA-1 and RNA-2 of CLRV I2. The RNA was probed with photobiotin-labelled pT7T3 18U clones 84 (lane 1), 57 (lane 2), 16 (lane 3), 22 (lane 4), 2 (lane 5), 53 (lane 6), 47 (lane 7), 25 (lane 8), 19 (lane 9), 3X (lane 10), 13 (lane 11) and 5 (lane 12).

(w/v) SDS for 30 min. The ends of each of these clones were sequenced; the pT7T3 18U-M13K07 single-stranded system (Pharmacia) was used to generate templates which were sequenced using a Sequenase kit (Sanger et al., 1977; dideoxynucleotide chain termination method).

This analysis allowed the clones to be designated as either RNA-1 clones (2, 13, 16, 19, 47, 53) or RNA-2 clones (5, 22, 25, 57, 84). A restriction map was produced that covered some 4 kb from the 3' ends of both RNA-1 and RNA-2 (Fig. 2). Several of the clones could not be fitted onto the map (clones 13, 25, 47, 53), presumably because their inserts were proximal to the 5' end of the virus RNAs.

The accuracy of the restriction map was confirmed after analysis of larger clones that covered this entire region. Clones 19 (RNA-1) and 5 (RNA-2) were used as probes (photobiotin-labelled) in the screening of a 2gtl0 (Promega) library [oligo(dT)12-18-primed]. Plaque lift filters were prepared and processed, and three additional RNA-1 clones of 3.5 kb (2gtl0; 1.1, 1.2 and 1.3) and two RNA-2 clones of similar size (2.1 and 2.2) were selected and found to have restriction sites that corresponded to those predicted by the map.

The entire sequences of the 3' end regions of RNA-1 and RNA-2 (1.8 kb) have been obtained by sequencing all the available clones that cover this region (Fig. 3). The sequencing showed that the two RNAs of I2 CLRV are virtually identical for over 1500 bases from the 3' end. The common sequence has some interesting features; it is polyadenylated, but no polyadenylation signal (AAUAAA, Proudfoot & Brownless, 1976; Fitzgerald & Shenk, 1981) was apparent. Also two U-rich regions were found at nucleotides 97, 228, 649, 1465 and 1482 in the RNA-1 sequence. The tRNA-like structure at nucleotide 97 (RNA-I) was absent from RNA-2, and owing to three base changes the structure at nucleotide 228 (RNA-1) was not formed by the RNA-2 sequence. It is not known whether these structures form in vivo or have any real function; however, it is interesting that the tRNA-like structure starting at nucleotide 228 is very close to the end of the RNA-1/RNA-2 duplication. Another feature of the sequence is that there are no open reading frames.

Fig. 2. Restriction maps of CLRV I2 cDNA clones and their positions relative to the 3' ends of RNA-1 and RNA-2. Hatched boxes show where sequence overlaps confirmed the position. Abbreviations used are: BamHI, B; EcoRI, E; HincII, H; HindIII, Hi; PstI, P; SacI, S; SpeI, Sp; SalI, Sa.
longer than 200 nucleotides, making it unlikely that the duplicated region codes for any large proteins. This observation confirms a previous study by Hellen et al. (1991) which suggested that the full potential coding capacity of the genome was not being used for the synthesis of the RNA-1 and RNA-2 polyproteins.

As the isolates of CLRV have similar genome sizes, we addressed the possibility that the duplicated region of the I2 isolate was also present in the genomes of other CLRV isolates. Two approaches were used to investigate this possibility. As the R and I2 isolates were the most different of the European CLRV isolates, we have cloned and sequenced several 3' end R clones. Using oligo(dT)12.18 as primer and unfractionated RNA as template, a cDNA library was made in pT7T3 18U. Four independent short (1 kb) clones (RI1, R25, R29 and R35) were identified, each with the same restriction enzyme map. When biotin-labelled dsDNA probes were made from two of these (R25 and R29) and used in Northern blot analysis (in parallel with I2 clone 3X), all three probes hybridized with both genomic strands of both the R and I2 isolates even after the 3' end terminal 250 bases had been removed. The explanation for the cross-reaction between the R and I2 isolates was found from the sequence analysis. The sequences of 1184 bases in clones R25 and R29 were identical and a run of approximately 700 nucleotides from the 3' end was very similar to the I2 sequence (Fig. 3).

The second approach used was to determine whether the 3' end probes discussed above hybridized with any of the other isolates. Three other isolates in particular were used for this study: a poplar isolate obtained from pollen of a poplar tree (Populus nigra) growing at Lechlade, Oxon., U.K., a beech isolate from Fagus sylvatica in Germany supplied by Dr A. T. Jones (Scottish Crop Research Institute, Invergowrie, U.K. and a walnut isolate from a Juglans regia growing in Oxford (Cooper & Edwards, 1980).
In a dot blot analysis (Nevins, 1987) I2 and R cDNA photobiotinylated probes hybridized well with the I2, R, beech, poplar and walnut CLRV isolates but not with RNA isolated from a lilac isolate of arabis mosaic nepovirus. These hybridization studies indicate that there is some sequence conservation across a range of isolates. Indeed the length of the 3' end sequence conservation between R and I2 implies importance for the sequence in the biology of CLRV. It is likely the terminal region has some role in replication/recognition but it is not clear what, if any, function the rest of the sequence may have. Database searches have revealed no major sequence identities to any plant or virus sequences. However, the sequence has short regions of similarity with plant and virus genes including TomRSV (70 bases) and with a birch pollen antigen (72% identity, 40 nucleotides). It is also not clear why the RNA-1, RNA-2 sequence may have. Database searches have revealed no important sequence conservation across a range of isolates.

References


(Received 22 July 1991; Accepted 18 October 1991)