Key words: PVY/potyviruses/nucleotide sequence

Molecular Cloning of a Potato Virus Y Genome: Nucleotide Sequence Homology in Non-coding Regions of Potyviruses

By TOM TURPEN
Biosource Genetics Corporation, 3333 Vaca Valley Parkway, Vacaville, California 95688, U.S.A.

(Accepted 29 March 1989)

SUMMARY
An aphid-transmissible field isolate of potato virus Y, strain MM, was purified from a pepper host and cloned. All but 18 terminal nucleotides of the 9.7 kb genome are apparently contained in two overlapping cDNA clones. The library of cDNA clones is likely to be representative of the viral RNA population present in infected plants because restriction endonuclease maps derived from cloned and uncloned cDNA are collinear, each region of the genome is represented by several independent clones, and nucleic acid sequencing of 5' and 3'-terminal regions revealed small AU-rich non-coding domains with blocks of nucleotide sequence homology between this strain of PVY and three other potyviruses. The efficient application of cDNA cloning techniques to a large positive-stranded RNA virus is described.

INTRODUCTION
Recent comparisons between homologous non-structural peptides of a few viruses indicate that the positive-stranded RNA viruses infecting plants may be related at a molecular level to only two supergroups both of which have counterparts in animal viral families (Goldbach, 1986). Potyviruses are placed in the 'picornavirus-like' supergroup. Their single genomic RNA molecule of approximately 9.5 kb is polyadenylated at the 3' end (Hari et al., 1979; Hellman et al., 1980), contains a small genome-linked virion protein (VPg) covalently linked at the 5' end (Hari, 1981; Siaw et al., 1985; Shahabuddin et al., 1988), and functions as messenger RNA coding for the synthesis of a single polypeptide precursor (Allison et al., 1986; Domier et al., 1986). This polyprotein (340K to 350K) is processed via an autocatalytic cascade (Carrington & Dougherty, 1987a,b; Hellmann et al., 1988) to a predicted seven or eight final peptides, including the coat protein which packages the RNA into flexuous rods 680 to 900 nm in length (Holling & Brunt, 1981). The most striking homology to mammalian picornaviruses is found in a cluster of genes encoding four polypeptides, all of which are thought to be involved in a membrane-bound replication complex (Domier et al., 1987). These proteins include a membrane-binding component, VPg, a protease and an RNA-dependent RNA polymerase, in conserved order.

Potyviruses constitute the largest and agronomically most important group of plant viruses known, causing devastating diseases in many crop plants (Hollings & Brunt, 1981). Potato virus Y (PVV) is the type member of the group, and strains of PVV cause diseases in tobacco, potato, pepper and tomato (De Boxx & Huttinga, 1981).

One approach to understanding the complex problems of RNA virus evolution and disease symptom induction requires the construction of cDNA clones of the complete genome of the virus and the ability to initiate infections from these clones. By recombining segments of viral genomes as DNA, phenotypes can easily be mapped physically to specific regions. In the cases of the RNA viruses, Qβ coliphage (Taniguchi et al., 1978), poliovirus (Racaniello & Baltimore, 1981; Semler et al., 1984; Omata et al., 1984) and coxsackie B3 virus (Kandolf & Hofschneider, 1985), the cDNA clones can be used to transfect cells directly. It is also possible to synthesize infectious transcripts, in vitro, from full-length cDNA clones of a number of RNA viruses,
including the plant-infecting bromoviruses (Ahlquist et al., 1984; Janda et al., 1987; Allison et al., 1988), tobacco mosaic virus (Dawson et al., 1986; Meshi et al., 1986) and cowpea mosaic virus (Vos et al., 1988), the human viruses rhinovirus (Mizutani & Colombo, 1985; Duchecler et al., 1989), poliovirus (Kaplan et al., 1985; Van der Werf et al., 1986) and Sindbis virus (Rice et al., 1987) and black beetle virus (Dasamahapatra et al., 1986). It would be useful to examine the sequence–function relationships of potyvirus components via recombinant DNA technology. Towards this goal, a library of large, overlapping cDNA clones was constructed from PVY virion RNA. Efficient cloning and sequencing techniques facilitated the rapid determination of the 5' and 3' sequence of the virus, and the verification of the clones produced.

METHODS

**Virion purification.** The isolate of PVY used in this work was collected 15 years ago from field-grown peppers in California by Dr. A. O. Paulus, Department of Plant Pathology, University of California, Riverside, Calif., U.S.A. The strain was subsequently propagated by Dr. L. G. Weathers (University of California, Riverside) on peppers and passaged yearly by single aphid transmission. In this work, the virus was propagated in mechanically inoculated peppers (Capsicum annuum cv. Mexican Chili), grown in greenhouses at an average temperature of 29 °C in the day (18h) and 24 °C at night (6h). Infected plants displayed a severe systemic mottling symptom upon infection. Extracts reacted positively to PVY antiserum (Agdia). Gooding (1985) derived a classification system capable of distinguishing nine strains of PVY, based on differential symptoms induced in a series of tobacco genotypes. The isolate (PVYmm) used in this work is an MM strain (mild mosaic), similar to the Y0 type strain (G. V. Gooding, personal communication; De Boks & Huttinga, 1981).

Virions were purified by homogenizing rinsed leaf tissue with an equal proportion (w/v) of phosphate buffer (50 mm-sodium phosphate pH 7.6, 10 mm-EDTA) and 0.5% (v/v) 2-mercaptoethanol in a blender for several minutes at a moderate speed (McDonald et al., 1976). After filtering through two layers of cheesecloth and Miracloth (Chicopee Mills), the homogenate was centrifuged at 12000 g for 30 min at 4 °C. The supernatant was layered over 15 ml of 30% (w/v) sucrose in phosphate buffer and centrifuged at 100000 g for 2.5 h at 4 °C (Beckman Ti-45 rotor). This crude virion pellet was resuspended in 4 ml of phosphate buffer and layered over 31% (w/w) CsCl in phosphate buffer and centrifuged at 120000 g for 20 h at 4 °C (Beckman SW-41 rotor). The virion band (ρ = 1.32 g/ml) was removed, mixed with the same CsCl solution and centrifuged again to equilibrium at 120000 g for 44 h at 4 °C (Beckman Ti-70 rotor). Virions were dialysed against 2 mm-sodium sulphite/phosphate buffer, negatively stained with 2% uranyl acetate, and examined by electron microscopy (Hitachi, H-600).

**RNA purification.** The RNA used in this study was isolated from purified virions by heating at 90 °C for 5 min in the presence of 1% SDS (2 mm-sodium sulphite/phosphate buffer). The RNA was concentrated by ethanol precipitation, resuspended in diethyl pyrocarbonate-treated water, 2% SDS (0.5 ml) and 100 to 200 μg was layered over 10 to 40% (w/v) sucrose gradients (RNase-free; Sigma) buffered in TNE (20 mm-Tris–HCl, 1 mm-NaCl, 5 mm-EDTA, pH 8-0). Gradients were centrifuged at 80000 g for 18 h at 20 °C (Beckman SW-28.1 rotor) and the RNA zone was recovered from the gradient and concentrated by ethanol precipitations. RNA length integrity was analysed by agarose gel electrophoresis of glyoxal-treated samples (McMaster & Carmichael, 1977).

**cDNA cloning.** Reaction conditions for cDNA cloning were those from the Amersham cDNA synthesis system based on the methods of Okayama & Berg (1982) and Gubler & Hoffman (1983) with the following modifications. The viral RNA template was denatured with methylmercuric hydroxide (Maniatis et al., 1982) and in some first strand synthesis reactions, Moloney murine leukemia virus (Bethesda Research Laboratories, BRL) or avian myeloblastosis virus (Life Sciences) reverse transcriptase and RNAsin (Promega Biotec) were substituted. All cloning reactions were primed with oligo(dT) during first strand synthesis and RNase H-digested template was the primer for second strand synthesis. Reactions were quantified by the incorporation of [α-32P]dCTP (New England Nuclear, NEN) in some first strand synthesis reactions, Moloney murine leukemia virus (Bethesda Research Laboratories, BRL) or avian myeloblastosis virus (Life Sciences) reverse transcriptase and RNasins (Promega Biotec) were substituted. All cloning reactions were primed with oligo(dT) during first strand synthesis and RNase H-digested template was the primer for second strand synthesis. Reactions were quantified by the incorporation of [α-32P]dCTP (New England Nuclear, NEN) and labelled with [γ-32P]ATP (NEN) by polynucleotide kinase (New England Biolabs), as M, standards. The double-stranded cDNA products were blunted-ended with T4 DNA polymerase and ligated to EcoRI or NotI linkers with T4 DNA ligase (IBI). In some experiments, internal restriction endonuclease digestion fragments were ligated into plasmids and cloned without linker addition. Restriction endonucleases were from NEB, BRL or IBI.

The cDNA was purified by sedimentation in sucrose gradients. The sucrose gradients were identical to those used in the fractionation of PVY RNA except that they were run at 100000 g for 20 h at 4 °C. A portion (20 μl) of each fraction was electrophoresed on agarose gels and the DNA transferred to a GeneScreen membrane (DuPont) by the alkaline blot method (Reed & Mann, 1983). Filters were hybridized with nick-translated clones as described previously (Turpen et al., 1987). When appropriate standards were included on the gel, this procedure served efficiently to fractionate the cDNA reaction products of different sizes, remove the linker trimmings, and quantify the amount of cDNA recovered for cloning ligations. DNA was concentrated from the desired fractions by ethanol precipitations.
The Bluescribe and Bluescript plasmids (Stratagene) and pBR322 (Bolivar et al., 1977) were used as cloning vectors following restriction enzyme digestion, phosphatase treatment and repurification by electroelution from agarose gels. *Escherichia coli* strains C600, HB101 and DH2 (BRL), JM109 and XLI Blue (Stratagene) were made competent for transformation by the procedure of Hanahan (1983).

**Nucleic acid sequencing.** Deletion subclones were constructed in Bluescript vectors using the exonuclease III/mung bean nuclease reactions and single-stranded template was prepared as directed by the supplier (Stratagene). The Sequenase (United States Biochemicals) dideoxynucleotide DNA sequencing protocol was used for reactions on single-stranded templates. For denatured double-stranded template, the Gem/Seq K/RT system was used (Promega Biotec). Synthetic DNA primers (Applied Biosystems) were prepared by Jim McKenzie (Zoeeon Research Institute).

The Gem/Seq system was used for dideoxynucleotide sequencing on the RNA (VPg-linked) template. Alternatively, some RNA sequencing and all of the primer extension reactions were done according to Geliebter et al. (1986). Chemicals were from Sigma or Bio-Rad. Nucleotides and dideoxynucleotides were from NEN. In some cases, the VPg-linked template (1 µg/µl) was pretreated with proteinase K (0.5 µg/0.1 µl), for 30 min at 37 °C, in 10 mM-Tris-HCl pH 7.5, 1 mM-EDTA.

Sequence data were analysed using programs from Intelligenetics and Textco.

**RESULTS**

A simple isolation procedure involving differential centrifugation of a crude homogenate, followed by two cycles of CsCl density gradient centrifugation (differential and then equilibrium) gave excellent yields of pure virions from pepper leaf tissue infected with the UCR strain of PVY. Purity was assessed by electron microscopy and particles were determined to be 734 ± 4 nm in length (mean ± s.e., n = 10; data not shown). Virion yields were 40 to 50 mg/kg tissue, and the ratio of absorbance (260 nm over 280 nm) was 1.2 to 1.3. Virions tended to aggregate irreversibly if pelleted in a pure form (Damirdagh & Shepherd, 1970).

The covalent linkage of a protein moiety to the 5' end of potyviral genomes (VPg) presents unusual difficulties in obtaining high quality RNA suitable for cloning. Standard extractions with organic solvents (Maniatis et al., 1982) yielded essentially no genome-length RNA, nor did CsCl/guanidinium thiocyanate centrifugation (Turpen & Griffith, 1986). Of several methods of RNA isolation from pure virions, RNA recovered from sucrose gradients was the most active in cDNA synthesis under standard cloning conditions. In the reaction shown in Fig. 1 (a) approx. 2 µg of double-stranded cDNA was synthesized from 5 µg of VPg-linked RNA.

The entire genome of PVY was cloned by the following five-step strategy. (i) Full-length, double-stranded cDNA was synthesized using oligo(dT) as a first strand primer and the final reaction products were mapped with restriction endonucleases (Fig. 1a, b). Cutting by the enzymes *KpnI*, *NotI*, *SmaI* or *Xhol* was not observed. (ii) Large cDNA clones were isolated using linkers containing restriction endonuclease sites not present in PVY cDNA (*NotI*). As expected, these clones proved to be 3'-coterminal and lacked significant portions of the 5' end of the genome (pZO171, pZO174; Fig. 1d). (iii) To clone the remaining portion of the genome, double-stranded cDNA was digested with two restriction endonucleases, one cutting at the terminal linker site (*NotI*) and one cutting at a site determined to be unique in PVY cDNA (*BamHI*) and also near the 5' end of the genome (Fig. 1c). The predicted 5'-terminal fragment (1-4 kb) was isolated from sucrose gradients and cloned into *NotI* and *BamHI*-digested vector. The sizes of seventy-five independent clones containing this fragment were determined at a resolution of 10 bp differences by agarose gel electrophoresis. The plasmid pZO177 and an apparently identical plasmid (pZO178) were found to contain the largest inserts. (iv) By similar techniques (see Methods) the library of clones was expanded such that each of the 10 *HindIII* fragments of the total genome (Fig. 1c) was present in several independent and overlapping clones. In addition to the 5'-terminal clones described above, the library contains 37 characterized clones having an average insert size of 3.4 kb. (v) The fine structure of the 5' and 3' termini was analysed by nucleic acid sequencing.

Six cDNA clones containing the 3' region of PVY were sequenced at the junction of the cDNA and vector. These inserts had poly(A) tracts of 1, 8, 8, 11 (pZO174) and 52 nucleotides, and pZO171 lacked three terminal nucleotides. No other sequence variation was detected between the independent clones analysed. The length variation at the 3' terminus does not
Fig. 1. (a) An autoradiograph (15 min exposure) of a dried agarose gel containing $^{32}$P-labelled cDNA (approx. 40 ng cDNA/lane), cut with restriction endonucleases HindIII, BamHI, PstI and SalI, and uncut (lanes 2 to 6, respectively). Lane 1 contains $^{32}$P-labelled Mr markers (1 kb ladder). Sizes are in kbp. (b) The relative fragment order and approximate sizes (given in parentheses) of uncloned cDNA were deduced from the restriction digestion analysis of products of varying size distribution from cDNA synthesis reactions. The orientation is 5' (VPg) to 3' [Poly(A) tail]. The complete PVY cDNA map (c) was compiled from the maps of several overlapping cDNA clones of which three are presented (d). Within the resolution of this analysis, the maps obtained with cloned and uncloned cDNA are corroborative. Restriction sites are as follows: (O), NotI; (●), HindIII; (□), PstI; (■), BamHI; (△), XhoI; (▲), SalI; (●), SacI + XhoI.
Fig. 2. On the basis of the DNA sequence of the 5' cDNA clones most distal to the poly(A) tail (pZO177 and pZO178), DNA primers were synthesized which hybridized at different positions near the 5' end of PVY RNA. The sequence of the remaining uncloned 15 nucleotides of the genome was determined by a combination of dideoxynucleotide sequencing and primer extension reactions using the VPg-linked genome as template. (a) The sequence obtained with a primer annealing 23 nucleotides from the end. The last cloned nucleotide present in both pZO177 and pZO178 is circled. (b) Regions of ambiguity at the terminus were resolved with the use of an additional set of synthetic DNA primers by removing individual nucleotides from the nucleotide mixture of primer extension reactions as indicated in the first four lanes of each panel. Lane 1 shows the template. Products in lane 2 are from reactions containing all four nucleotides. Pretreatment of the template with proteinase K (lane 3) to remove the VPg had no detectable effect on the extension or sequencing reactions.

The nucleotide sequence of pZO177 and pZO178 was used to synthesize DNA primers for dideoxynucleotide sequencing of the RNA template. This analysis indicated that a 15 nucleotide long viral sequence was missing from the corresponding 5' end region of the cDNA clones. A primer was radiolabelled with \(\gamma^{32}\)PATP, annealed 23 nucleotides from the end of the PVY genome and extended with reverse transcriptase in the presence of dideoxynucleotides (Fig. 2a). Further experiments were necessary to confirm the precise sequence of the last 15 nucleotides, particularly at the -6 position and at the terminus. Additional primers were synthesized to anneal at different positions near the 5' end of the genome (Fig. 2b). Either A, C, G or T were singly omitted from primer extension reactions containing the other three nucleotides (no dideoxynucleotides present). No extension was observed in the absence of template, and polymerization was either not initiated or terminated when the complementary nucleotide was not present in the reaction mix, enabling the sequence to be fully determined. Pretreatment of template with proteinase K to remove VPg had no detectable effect on sequencing or primer extension reactions (Fig. 2b).

The 5' and 3' non-coding regions of PVYmm are presented in Fig. 3. A set of deletion subclones of pZO177 and pZO171 have been constructed for the complete sequence analysis of the genome (to be published elsewhere). Translation is predicted to initiate at the first AUG codon at position 168. This region is in agreement with the plant consensus sequence for
Fig. 3. The nucleotide sequence of the non-coding region of PVYmm determined from viral RNA and from the cDNA clones pZO177, pZO171 and pZO174. The symbol < indicates the position of the last 3' polynucleotide of six cDNA oligomers used in sequencing the RNA template. Primers used for DNA sequencing annealed to both strands outside the region shown. The 5' end of the RNA genome contains the presumptive site of covalent linkage to the VPg. Ac. ) is the site of a variable length poly(A) tract at the 3' end. The clone pZO177 lacks the Y-terminal 15 nucleotides and pZO171 is missing the T-terminal three nucleotides and poly(A) tail. An obvious potential initiation region is found at position 168, with the plant consensus sequence for this region indicated below (Lutcke et al., 1987). Partial sequencing data indicate that the reading frame initiated at this first potential start codon from the 5' end extends for at least 1400 bases and the reading frame immediately preceding the stop codon at position 269 extends for at least 700 bases and includes the coat protein gene of PVY (Shukla et al., 1986; data not shown).

translation initiation sites, notably at the important + 4, + 5 and −3 positions. At the 3' end, the stop codon of the coat protein gene has been identified at position −269 by comparison to a published PVY coat protein sequence (Shukla et al., 1986). The 5' and 3' non-coding regions contain 72% and 61% A + U nucleotides, respectively.

The nucleotide sequence of both 5'- and 3'-terminal non-coding regions has been established for three other potyviruses, i.e. tobacco etch virus (TEV; Allison et al., 1986), tobacco vein mottling virus (TVMV; Domier et al., 1986) and the N strain of potato virus Y (PVYN; Robaglia et al., 1989), enabling the identification of common sequences, secondary structures and repeats of possible functional significance to the potyvirus group by comparative analysis.

The 3' non-coding regions of each virus sequence can be folded (Zuker & Steigler, 1981) into stable secondary structures (PVYn, −20 kJ; PVYmm, −16 kJ; TVMV, −13.6 kJ; TEV, −11 kJ). Similarly, several hairpin loops (Tinoco et al., 1973) can be identified. Some of the more stable hairpin loops including those previously identified for TVMV (Domier et al., 1986) are not present in the most stable foldings of the entire region. Although structures common to the four viruses may exist, they are not readily apparent from this analysis. Primary sequence conservation in this region is limited between members of the group to the sequence AUAAUAUA found within 60 nucleotides following the stop codon of the polyprotein reading frame of each virus. This sequence is often present in plant mRNAs (Joshi, 1987). The organization of a direct repeat element, first identified by comparison between another isolate of
PVY cloning and potyvirus sequence homology

Fig. 4. Boxes 'a' and 'b' were identified as regions containing the highest number of consecutive nucleotides with sequence identity in the non-coding regions of all four viruses. Nucleotides are numbered from the 5' end.

<table>
<thead>
<tr>
<th>Virus</th>
<th>Box a</th>
<th>Box b</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVMV</td>
<td>13 ucaAcACAACAUua 26</td>
<td>39 aaucUCAAGCAaucaagc 56</td>
</tr>
<tr>
<td>TEV</td>
<td>13 ucaAcACAACAUau 26</td>
<td>39 aaucUCAAGCAaucaagc 56</td>
</tr>
<tr>
<td>PVYn</td>
<td>13 ucaAuACAACAUaa 26</td>
<td>63 ucaucUCAAGCAacuugcu 80</td>
</tr>
<tr>
<td>PVYmm</td>
<td>13 aacAuACAACAUaa 26</td>
<td>41 auuuUCAAGCAuuuuucu 58</td>
</tr>
</tbody>
</table>

PVYn and pepper mottle virus (PeMV; van der Vlugt et al., 1989) is weakly discernible in the 3' region of PVYmm and is present as a 32 nucleotide perfect direct repeat in the PVYn isolate sequenced by Robaglia et al. (1989). Within this repeat a core 6 nucleotide sequence, GTGACT, is present in all PVY strains but absent from TVMV and TEV.

In contrast, two sets of 7 residue sequences (Fig. 4) were found in the 5' untranslated regions of all four viruses. The 5' non-coding regions were folded into comparatively less stable structures than the 3' non-coding regions and no common structures were observed between the three. Upon closer examination, it is apparent that sequences related to the major blocks of homology are repeated throughout the regions. On an overlapping basis, 'box a' is repeated a total of eight times with a mismatch of one nucleotide, and 'box b' is repeated nine times with a mismatch of one nucleotide.

**DISCUSSION**

Currently, there is great interest in the cDNA cloning of RNA viruses. Advances in recombinant DNA technology have suggested numerous experiments with important implications in basic and applied biology of the RNA viruses and their hosts, particularly in plant systems where the host organism can be easily genetically transformed. However, many technical cloning problems are accentuated in cloning large RNAs.

The first problem is to distinguish cloning artefacts from authentic cDNA. For example, the ligation of two cDNA fragments into a vector instead of one, or the failure to recover small internal restriction fragments, would complicate subsequent analysis. Secondly, smaller changes may affect biological activity of the cDNA and may be derived either from sequence variation in the original RNA population, an error in reverse transcription of the first strand or mismatch repair of random primers. The final obstacle in cloning an uncharacterized genome is to obtain terminal sequence information of good quality rapidly, particularly for the 5' end. Once these sequences are known, several other cloning strategies might be adapted. The strategy used in the research reported here on PVY addresses each of these problems.

To ensure that cloned cDNA was representative of virion RNA, the cDNA was mapped with restriction endonucleases before cloning and compared to cloned cDNA. To overcome problems of sequence variation, many independent, overlapping clones of each region were obtained. Oligo(dT) was the primer for all first strand cDNA synthesis used in cloning experiments. Essentially, a one-step cloning 'walk' was used to isolate and clone a predicted 5'-terminal fragment. Enough independent clones of the 5' region were examined to ensure few nucleotides would be missing, such that the uncloned region could be easily resolved on a single sequencing gel.

Dideoxynucleotide sequencing and primer extension reactions on the RNA template resolved the remaining sequence of 15 terminal nucleotides missing from pZO177 and pZO1778 (Fig. 2). The presence of the VPg did not appear to interfere with reverse transcriptase-catalysed synthesis of the cDNA strand (Fig. 2b). Similar results were obtained in sequencing of the 5' termini of human rhinovirus (Callahan et al., 1985) and TVMV (Shahabuddin et al., 1988). An additional faint T nucleotide was detected in the sequencing reaction in Fig. 2(a) but not in the primer extension reactions in Fig. 2(b). Also, there were some strong bands below the final T particularly at the −6 position. It remains to be shown whether this is due to premature termination or length heterogeneity as it has been suggested for TVMV (Shahabuddin et al., 1988).
The PVY sequence presented is correct for the majority of the RNA population, and the missing nucleotides have been replaced in the cDNA clones by oligonucleotide-directed repair (data not shown). The size of the resulting cDNA restriction map of PVY is in good agreement with the size of the RNA genome.

This cloning strategy might be further improved. For the level of competence of the cells used in the transformations and the amounts of DNA used under a variety of optimized ligation conditions, a small number of transformants was obtained (see also Kandolf & Hofschneider, 1985). Also, the size distribution of inserts cloned was significantly smaller than the size distribution of the cDNA molecules synthesized and purified from sucrose gradients. Perhaps for large cDNAs, such as those synthesized from potyviral RNA, ligation to phage instead of plasmid cloning vehicles would yield better results.

The blocks of primary sequence homology identified in the 5' non-coding regions are likely to be functionally significant because they are conserved in each of four divergent potyvirus isolates. Pairwise computer alignments between each virus (five comparisons) reveal a maximum of only 58 to 65% homology for this region (data not shown). Blocks of 5' non-coding region homology are also reported in members of the family Picornaviridae (Skern et al., 1987), where the region is implicated in virulence and translation functions (Pilipenko et al., 1989). An experimental approach will be necessary to define the structure and function of these regions.

Primary sequence is even less conserved at the 3' end (55 to 59% alignments, data not shown). These comparisons provide further support for the suggestion that PeMV is best classified as a strain of PVY because the homology is far greater in this region between PeMV and some PVY strains than the homology found between PVY strains (Shukla & Ward, 1988; van der Vlugt et al., 1989).

Once the prerequisite cDNA clones have been obtained, it should be possible to use genetic engineering techniques to introduce viral disease resistance into transgenic plants by a number of mechanisms (Sanford & Johnston, 1985). This approach was recently reviewed in the context of inactivating cellular genes (Herskowitz, 1987). The challenge is to design genes that have activity towards several or all members of a group of viruses. The problems of viral control and viral evolution are inseparable. Effective antiviral genes must not only inhibit an essential step in the viral replication cycle and/or symptom-inducing function but must also target a region of the genome where future sequence variation will result in reduced fitness and virulence of the pathogen. The common sequences identified in this work are large enough to provide a reasonable target for engineering sequence-specific ribozyme activity against potyviruses (Haseloff & Gerlach, 1988).

I am grateful to many colleagues at Zoeeon Research Institute, Sandoz Crop Protection Corp., and the University of California at Riverside where most of this work was completed, for their encouragement and support, including Jim Young, Larry Grill, Bill Dawson, Steve Garger, Ann Myers, Kathy Konzak and Allan Bracker. Dr G. V. Gooding, Jr (North Carolina State University) provided valuable advice on PVY symptomatology and Dr C. Robaglia (INRA) shared the sequence of PVYn before publication.

REFERENCES


(Received 16 November 1988)