J. gen. Virol. (1989), **70**, 783–788. *Printed in Great Britain*

Key words: *PLRV/coat protein/nucleotide sequence*

Identification and Characterization of the Potato Leafroll Virus Putative Coat Protein Gene

By L. M. KAWCHUK,[1]* R. R. MARTIN,[1] D. M. ROCHON[1] AND J. McPHERSON[2]

[1]*Research Station, Agriculture Canada, 6660 North-West Marine Drive, Vancouver, British Columbia V6T 1X2 and* [2]*Department of Plant Science, 2357 Main Mall, MacMillan Building, Room 341, University of British Columbia, Vancouver, British Columbia V6T 2A2, Canada*

(*Accepted 9 November 1988*)

SUMMARY

Complementary DNA clones representing approximately 6100 nucleotides of potato leafroll virus (PLRV) were generated, restriction-mapped, and partially sequenced. Within one of the cDNA clones an open reading frame (ORF) encoding a 23K protein was identified and further characterized. Amino acid sequence comparison of this protein showed significant homology (47·1%) with the barley yellow dwarf virus (BYDV-PAV) coat protein. This and other observations suggested that this gene encodes the PLRV coat protein. Other similarities were observed between PLRV and BYDV sequences in this region of their genomes, including an ORF of 17K within the ORF encoding the 23K putative coat protein.

Potato leafroll virus (PLRV), a member of the luteovirus group (Matthews, 1982) is difficult to control and is responsible for significant economic losses in potato (*Solanum tuberosum* L.) (Rowhani & Stace-Smith, 1979). It is a phloem-limited spherical virus transmitted by several aphid species in a persistent manner, the most efficient vector being *Myzus persicae* Sulz. (Harrison, 1984). Virions of PLRV are composed of a single-stranded RNA species of M_r $2{\cdot}0 \times 10^6$ and a single coat protein of M_r 26·3K (Rowhani & Stace-Smith, 1979).

The coat protein (CP) of luteoviruses is reported to be responsible for many viral characteristics including serological properties (Waterhouse & Murant, 1981), transmission specificity (Massalski & Harrison, 1987; Rochow & Carmichael, 1979), cross-protection (Harrison, 1958; Webb *et al.*, 1952), and mutual exclusion (Jedlinski & Brown, 1965). The coat protein genes of several viruses have been expressed in transgenic plants resulting in reduced virus synthesis following specific exposure to virus (Abel *et al.*, 1986; Cuozzo *et al.*, 1988; Hemenway *et al.*, 1988; Loesch-Fries *et al.*, 1987). The identification, characterization, and comparison of PLRV and other luteovirus coat proteins should permit an understanding of their specific roles.

An isolate of the PLRV characterized by Rowhani & Stace-Smith (1979) was purified by a modification of their procedure. The yield of virus was increased by incubating the plant homogenate in 0·1% (w/v) Ultrazym 100 (Schweizerische Ferment) overnight at room temperature. RNA was purified from virus particles by vortex-mixing in 0·38 M-Tris–HCl pH 8·9, 20 mM-EDTA, 1% (w/v) SDS followed by phenol–chloroform extraction.

Double-stranded cDNA copies of PLRV RNA were synthesized and cloned as follows. Virion RNA, polyadenylated *in vitro* by the procedure of Sippel (1973), served as template for oligo(dT)-primed first strand cDNA synthesis. Single-stranded cDNA was also prepared using virion RNA as template for randomly primed cDNA synthesis (Maniatis *et al.*, 1982). Second strand DNA synthesis was conducted using DNA polymerase I and RNase H-generated primers (Gubler & Hoffman, 1983). Double-stranded cDNA was dC-tailed using terminal

0000-8678 © 1989 SGM

transferase (Bethesda Research Laboratories, BRL) and then annealed with *Pst*I-restricted pUC9 similarly tailed with dG. Half of the ds cDNA from the randomly primed reaction was treated with mung bean nuclease and then ligated into *Eco*RV-restricted Bluescript M13 vector (Stratagene) which had been treated with calf intestinal phosphatase. Following ligation the DNA was used to transform competent DH5α cells (BRL) which were plated on Luria–Bertani agar containing ampicillin and X-gal. Two-hundred white, ampicillin-resistant colonies obtained by the oligo(dT) method and 800 colonies by the random priming method were screened by colony filter hybridization (Gergen *et al.*, 1979) using randomly primed ^{32}P-labelled PLRV cDNA probes. Those colonies giving the strongest signal (approx. 100) were selected and recombinant plasmids were isolated using the alkaline lysis method (Maniatis *et al.*, 1982). Plasmids were analysed by digestion with *Pst*I or *Pvu*II, which cleaved out the inserts, and were then electrophoresed in agarose gels. Those colonies containing the plasmids with the largest cDNA inserts were selected for further study.

Potato leafroll virus RNA and total RNA from healthy potato leaves were electrophoresed on denaturing methylmercury gels and transferred to nitrocellulose (Schleicher & Schuell) as described by Maniatis *et al.* (1982). The origin of clones B2a, LP79 and LP93 (Fig. 1*a*) was confirmed by hybridization to PLRV RNA using Northern blot analysis. ^{32}P-labelled probes were synthesized from each cDNA clone using the randomly primed oligonucleotide method (Feinberg & Vogelstein, 1983). Each probe hybridized specifically to PLRV RNA and was negative with uninfected plant RNA (data not shown).

The clones shown in Fig. 1 (*a*) were oriented relative to each other based on restriction enzyme and hybridization analyses. For hybridization, end fragments were gel-purified, oligonucleotide-labelled, and used to select overlapping clones. Plasmid DNA was spotted onto Nytran (Schleicher & Schuell) sheets and subsequently denatured by placing the Nytran on a stack of 3MM filter paper soaked in 0·5 M-NaOH.

Five clones were selected and found to cover approx. 6100 nucleotides of the virus genome which is greater than 98% (Fig. 1*a*) of the predicted size of PLRV genomic RNA (6200 nucleotides) as determined by methylmercury gel electrophoresis. Clone B2a and several other oligo(dT)-primed clones mapped to the same location suggesting that they were derived from the 3′ end of PLRV RNA. Although some differences were observed, restriction analysis revealed overall similarities with the restriction sites of the PLRV isolates described by both Prill *et al.* (1988) and Smith *et al.* (1988).

Recently the complete nucleotide sequence of the luteovirus barley yellow dwarf (BYDV-PAV) was determined (Miller *et al.*, 1988*a*) and the CP gene was shown to be located internally. If the PLRV CP were similarly located the clone LP79 (Fig. 1*a*) could contain the CP coding region of PLRV. This clone covered a large proportion of the internal region of the PLRV genome and therefore was subcloned and sequenced in both orientations (Fig. 1*b*).

Clones to be sequenced were inserted into the Bluescript vector and unidirectional nested deletions were prepared using exonuclease III and mung bean nuclease (Henikoff, 1984). Subclones were sequenced by the dideoxynucleotide chain termination method of Sanger *et al.* (1977) using double-stranded plasmid DNA as template (Korneluk *et al.*, 1985). Sequences were read using a Bio-Rad digitizer and analysed using Bio-Rad Gene-Master software.

The nucleotide sequence of LP79 revealed an open reading frame (ORF), nucleotides 34 to 657, as shown in Fig. 2. The predicted amino acid sequence (Fig. 2) shows extensive homology of 47·1% (Fig. 3*a*) with the amino acid sequence of the CP of BYDV (Miller *et al.*, 1988*b*). The predicted M_r of the PLRV ORF protein of 208 amino acids is 23 202 (23K) which is similar to the previously estimated M_r of 26·3K for the PLRV CP (Rowhani & Stace-Smith, 1979). The above evidence strongly suggests that the 23K protein encoded by the 23K ORF is the coat protein of PLRV. Like BYDV CP the amino-terminal region of the PLRV 23K ORF is highly basic, containing many arginine residues (Fig. 2). The amino-terminal regions of many plant virus capsid proteins have been found to be highly basic and it has been suggested that these regions may be involved in protein–RNA interaction (Harrison, 1983). Comparison of the PLRV 23K ORF nucleotide sequence with that of the BYDV CP gave 58·0% similarity. This reflects the amino acid similarity and indicates a close evolutionary relationship between these two viruses.

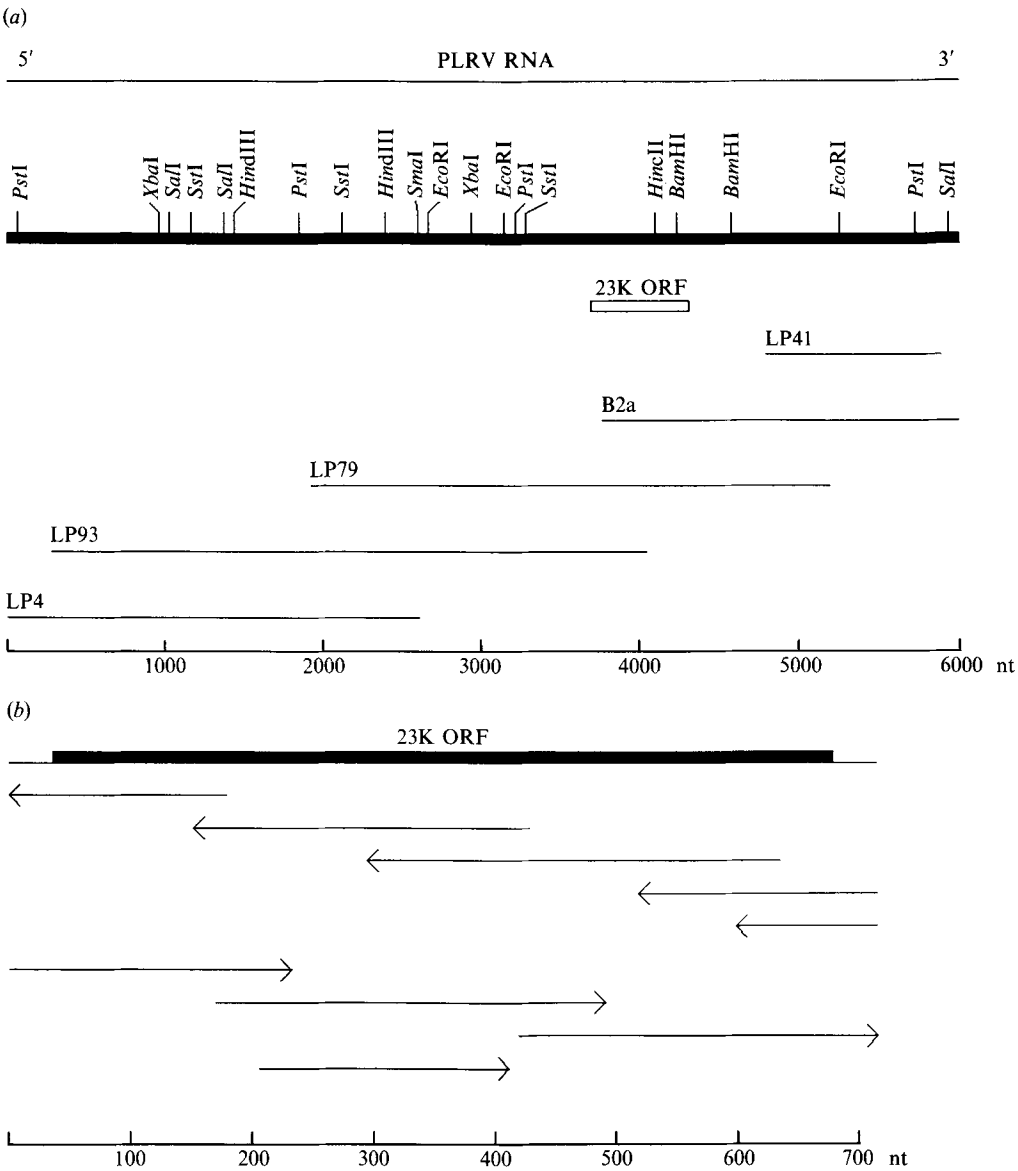

Fig. 1. (*a*) Restriction map of the PLRV genome derived from the analysis of overlapping cDNA clones that are indicated below the map. Although not shown, each *Sal*I site was also restricted by *Hin*cII. Clone B2a was from oligo(dT)-primed reactions and those of the LP series from the randomly primed reactions. The region of the clone LP79 corresponding to the 23K ORF is represented by an open rectangle. The line above the map corresponds to the PLRV RNA. (*b*) Sequencing strategy used to determine the putative coat protein coding region of PLRV. Arrows indicate the direction and length of sequence from unidirectional nested deletions of clone LP79.

Within the 23K ORF coding region but in a different reading frame lies another ORF, nucleotides 59 to 526, coding for 156 amino acids. Translation of this ORF would yield a protein of M_r 17381. BYDV and soybean dwarf virus (Miller *et al.*, 1988*b*) both contain a 17K ORF within the region encoding the CP. The similar finding of the 17K ORF within the PLRV putative CP gene suggests that this protein is a functional gene product. Miller *et al.* (1988*b*) suggested that the BYDV 17K protein is the genome-linked viral protein (VPg) based on its size

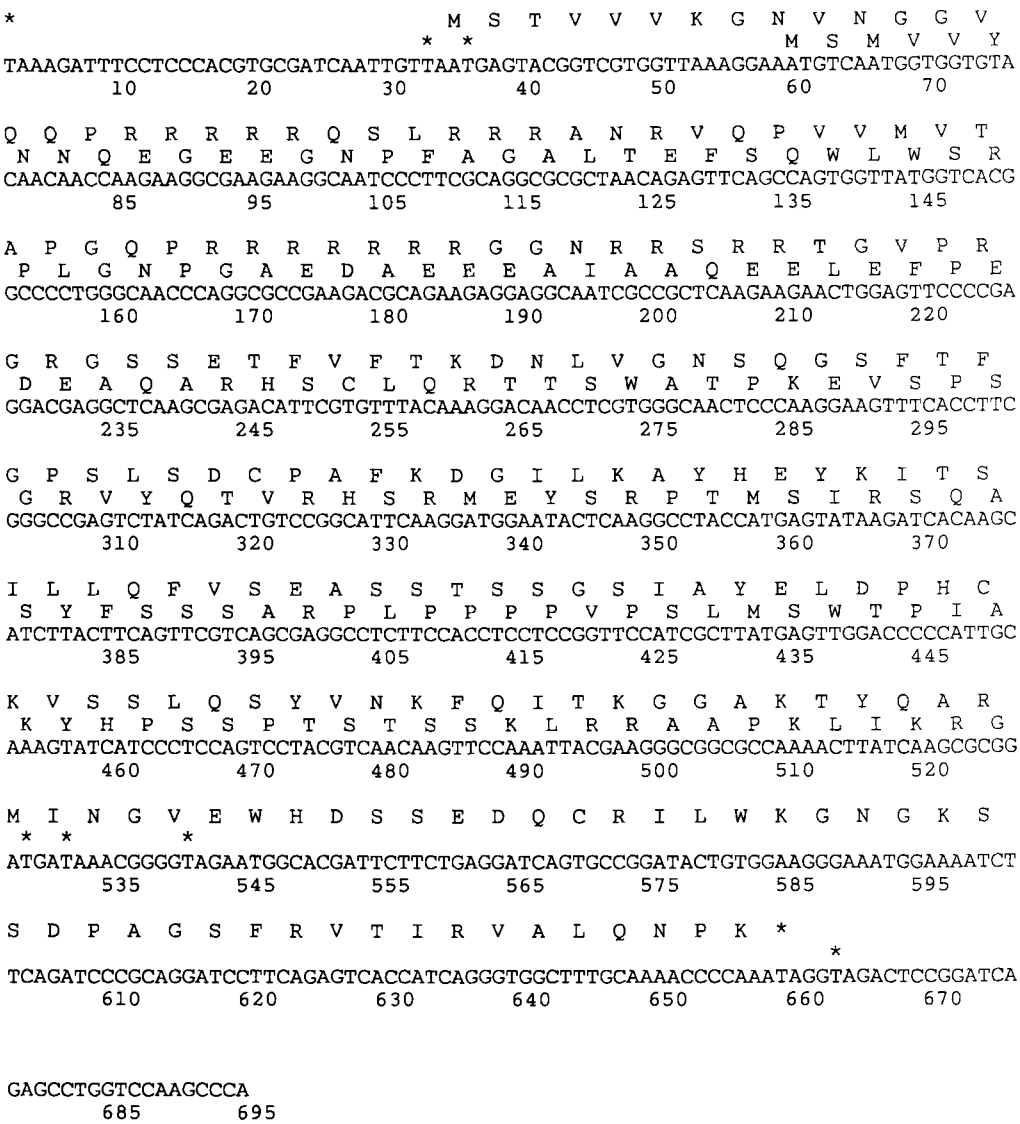

Fig. 2. Nucleotide sequence and predicted amino acid sequence of the PLRV 23K ORF and 17K ORF. The start codons used for the 23K ORF and 17K ORF are the first AUG codons following a stop codon (★) in the same frame.

being similar to isolated VPg. The PLRV VPg isolated by Mayo *et al.* (1982) has been estimated as 7K. There is less homology (31·1%) between the 17K ORFs of PLRV and BYDV (Fig. 3*b*) than there is between the putative coat proteins. Thus, if a functional protein is encoded by ORF 17K, it may be responsible for some shared but distinct property of each virus such as host specificity. Other examples of plant RNA viruses which indicate extensive use of overlapping ORFs are southern bean mosaic virus (Wu *et al.*, 1987) and carnation mottle virus (Guilley *et al.*, 1985).

The 23K ORF in PLRV terminates with an amber codon (Fig. 2) which is immediately followed by a long in-frame ORF (L. M. Kawchuk *et al.*, unpublished data). A similar finding was observed by Miller *et al.* (1988*b*) for BYDV and it was suggested by these authors that

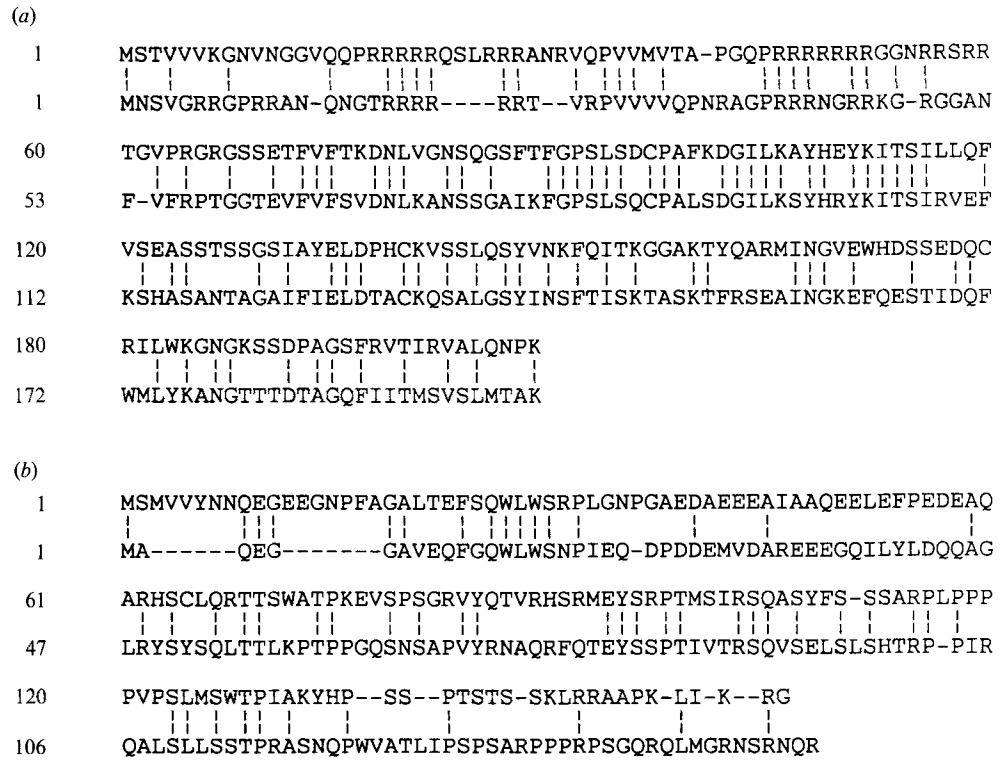

Fig. 3. (*a*) An alignment of the amino acid sequence of the PLRV 23K ORF (upper) and BYDV coat protein (lower). (*b*) An alignment of the amino acid sequences of the PLRV 17K ORF (upper) and the BYDV 17K protein (lower). The Needleman–Wunsch algorithm (Needleman & Wunsch, 1970) was used to align the sequences.

readthrough of the amber codon may occur. In both cases several proline codons occur in the region immediately following the amber codon; for PLRV seven of 13 amino acids in this region are proline (data not shown).

Although the similarity of the predicted amino acid sequence of the PLRV ORF 23K and BYDV proteins was high (47·1%), when intact virus was used as antigen and immunogen there was a lack of serological cross-reactivity. However, extensive cross-reactivity was observed when denatured virus was used as antigen (R. R. Martin & C. J. D'Arcy, unpublished data). This suggests that the most variable regions of the luteovirus capsid are exposed in intact virions.

The isolation and expression of the PLRV CP gene will provide sufficient amounts of protein for the production of large quantities of sequence-specific antibodies. These are presently only produced in small amounts because of difficulties in obtaining large amounts of viral coat protein. Comparison of luteovirus coat proteins will also allow the prediction of common antigens, i.e. regions which may be exploited in the preparation of antibodies that react with several if not all luteoviruses. Strain-specific epitopes may also be predicted by expressing small regions of the CP gene, producing antibodies and examining their specificity. Finally, since earlier investigators demonstrated that different strain combinations of luteoviruses give cross-protection or mutual exclusion, it should be possible to construct chimeric genes that when inserted into an appropriate host may permit further examination of these phenomena.

We wish to thank Dr W. A. Miller for an advance copy of the BYDV coat protein sequence and Dr B. Prill for a copy of the submitted manuscript on the cloning of PLRV. L.M.K. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada via a postgraduate scholarship.

REFERENCES

ABEL, P. P., NELSON, R. S., DE, B., HOFFMANN, N., ROGERS, S. G., FRALEY, R. T. & BEACHY, R. N. (1986). Delay of disease development in transgenic plants that express the tobacco mosaic virus coat protein gene. *Science* **232**, 738–743.

CUOZZO, M., O'CONNELL, K. M., KANIEWSKI, W., FANG, R., CHUA, H. & TUMER, N. E. (1988). Viral protection in transgenic tobacco plants expressing the cucumber mosaic coat protein or its antisense RNA. *Bio/Technology* **6**, 549–557.

FEINBERG, A. P. & VOGELSTEIN, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochemistry* **132**, 6–13.

GERGEN, J. P., STERN, R. H. & WENSINK, P. C. (1979). Filter replicas and permanent collections of recombinant DNA plasmids. *Nucleic Acids Research* **7**, 2115–2136.

GUBLER, U. & HOFFMAN, B. J. (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263–269.

GUILLEY, H., CARRINGTON, J. C., BALAZS, E., JONARD, G., RICHARDS, K. & MORRIS, T. J. (1985). Nucleotide sequence and genome organization of carnation mottle virus RNA. *Nucleic Acids Research* **13**, 6663–6677.

HARRISON, B. D. (1958). Ability of single aphids to transmit both avirulent and virulent strains of potato leaf roll virus. *Virology* **6**, 278–286.

HARRISON, B. D. (1984). Potato leafroll virus. *CMI/AAB Descriptions of Plant Viruses*, no. 291.

HARRISON, S. C. (1983). Virus structure: high resolution perspectives. *Advances in Virus Research* **28**, 175–240.

HEMENWAY, C., FANG, R., KANIEWSKI, W. K., CHUA, N. & TUMER, N. E. (1988). Analysis of the mechanism of protection in transgenic plants expressing the potato virus X coat protein or its antisense RNA. *EMBO Journal* **7**, 1273–1280.

HENIKOFF, S. (1984). Unidirectional digestion with exonuclease III creates target breakpoints for DNA sequencing. *Gene* **28**, 351–359.

JEDLINSKI, H. & BROWN, C. M. (1965). Cross protection and mutual exclusion by three strains of barley yellow dwarf virus in *Avena sativa* L. *Virology* **26**, 613–621.

KORNELUK, R. G., QUAN, F. & GRAVEL, R. A. (1985). Rapid and reliable dideoxy sequencing of double-stranded DNA. *Gene* **40**, 317–323.

LOESCH-FRIES, L. S., MERLO, D., ZINNEN, T., BURHOP, L., HILL, K., KRAHN, K., JARVIS, N., NELSON, S. & HALK, E. (1987). Expression of alfalfa mosaic virus RNA4 in transgenic plants confers virus resistance. *EMBO Journal* **6**, 1845–1851.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: A Laboratory Manual.* New York: Cold Spring Harbor Laboratory.

MASSALSKI, P. R. & HARRISON, B. D. (1987). Properties of monoclonal antibodies to potato leafroll luteovirus and their use to distinguish virus isolates differing in aphid transmissibility. *Journal of General Virology* **68**, 1813–1821.

MATTHEWS, R. E. F. (1982). Classification and nomenclature of viruses. *Intervirology* **17**, 1–199.

MAYO, M. A., BARKER, H., ROBINSON, D. J., TAMADA, T. & HARRISON, B. D. (1982). Evidence that potato leafroll virus RNA is positive-stranded, is linked to a small protein and does not contain polyadenylate. *Journal of General Virology* **59**, 163–167.

MILLER, W. A., WATERHOUSE, P. M. & GERLACH, W. L. (1988*a*). Sequence and organization of barley yellow dwarf virus genomic RNA. *Nucleic Acids Research* **16**, 6097–6111.

MILLER, W. A., WATERHOUSE, P. M., KORTT, A. A. & GERLACH, W. L. (1988*b*). Sequence and identification of the barley yellow dwarf virus coat protein gene. *Virology* **165**, 306–309.

NEEDLEMAN, S. B. & WUNSCH, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology* **48**, 443–453.

PRILL, B., MAISS, E., CHANSILPA, N. & CASPER, R. (1988). Molecular cloning of single-stranded RNAs of potato leafroll virus and beet western yellows virus. *Journal of General Virology* **69**, 2397–2402.

ROCHOW, W. F. & CARMICHAEL, L. E. (1979). Specificity among barley yellow dwarf viruses in enzyme immunosorbent assays. *Virology* **95**, 415–420.

ROWHANI, A. & STACE-SMITH, R. (1979). Purification and characterization of potato leafroll virus. *Virology* **98**, 45–54.

SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences, U.S.A.* **74**, 5463–5467.

SIPPEL, A. E. (1973). Purification and characterization of adenosine triphosphate : ribonucleic adenyltransferase from *Escherichia coli*. *European Journal of Biochemistry* **37**, 31–40.

SMITH, O. P., HARRIS, K. F., TOLER, R. W. & SUMMERS, M. D. (1988). Molecular cloning of potato leaf roll virus complementary DNA. *Phytopathology* **78**, 1060–1066.

WATERHOUSE, P. M. & MURANT, A. F. (1981). Purification of carrot red leaf virus and evidence from four serological tests for its relationship to luteoviruses. *Annals of Applied Biology* **97**, 191–204.

WEBB, R. E., LARSON, R. H. & WALKER, J. C. (1952). Relationships of potato leaf roll virus strains. *Research Bulletin. Agricultural Experimental Station, College of Agriculture, University of Wisconsin* **178**, 1–38.

WU, S., RINEHART, C. A. & KAESBERG, P. (1987). Sequence and organization of southern bean mosaic virus genomic RNA. *Virology* **161**, 73–80.

(Received 20 September 1988)