NUCLEOTIDE SEQUENCE ANALYSIS OF THE ORF0 REGION OF POTATO LEAFROLL VIRUS ISOLATES FROM TUNISIA

F. DJILANI-KHOUADJA[1*], J. ROUZÉ-JOUAN[2], M. MARRAKCHI[1], H. FAKHFAKH[1*]

[1]Laboratoire de Génétique Moléculaire, Immunologie et Biotechnologie, Faculté des Sciences de Tunis, Campus Universitaire El Manar, 2092 Tunis, Tunisie; [2]Unité Mixte de Recherche INRA/ENSAR, Biologie des Organismes et des Populations Appliquées à la Protection des Plantes, Rennes, France

Received November 30, 2005; accepted April 7, 2006

Summary. – Potato leafroll virus (PLRV) isolates from potato plants from different regions of Tunisia were investigated for the ORF0 variable region of genomic RNA using PCR, nucleotide sequencing and sequence analysis. Based on the ORF0 variable region nucleotide sequence, individual Tunisian isolates were more homologous as a group compared to the isolates originating from other parts of the world. There was no correlation between bioclimatic origin of the isolates and their ORF0 sequence. Alignment of the deduced amino acid sequences showed that the P0 protein was not much conserved. Unexpectedly, Tunisian isolates were found to be most homologous to Peruvian ones both at nucleotide and amino acid level. A phylogenetic tree, based on the P0 amino acid sequence, showed that all the PLRV isolates were located in two major clusters regardless of their geographic origin. In the second cluster, three sub-clusters could be distinguished. These results provide valuable information for molecular characterization of the PLRV isolates occurring in Tunisia.

Key words: ORF0; nucleotide sequence; phylogenetic analysis; Potato leafroll virus; Tunisian isolates

Introduction

PLRV (the species *Potato leafroll virus*, the genus *Polerovirus*, the family *Luteoviridae*) is responsible for significant worldwide economic losses in potato crops (Robert and Lemaire, 1999). It is transmitted in a persistent manner by some aphid species (Waterhouse *et al*., 1987) including *Myzus persicae*, the most efficient vector (Harrison, 1984). The natural PLRV host range is mainly restricted to a few solanaceous plants including *Physalis floridana* (Rydb) and *Solanum tuberosum* (Thomas, 1993). Small isometric virus particles of 25–30 nm in diameter are largely confined to the phloem and companion cells of infected plants (Shepardson *et al*., 1980).

PLRV genome consists on a single-stranded positive-sense RNA comprising eight ORFs (Ashoub *et al*., 1998). It is densely packed, with ORF1 overlapping much of ORF0, and ORF4 being entirely within ORF3. The three 5'-proximal ORFs (ORF0-ORF2) are expressed from the genomic RNA and the five 3'-proximal ORFs (ORF3-ORF7) are expressed from subgenomic mRNAs that are 3'-coterminal with genomic RNA. Thus luteoviruses represent a good example of the versatility of expression strategy and the economy of using the coding capacity of small RNA virus genomes (Taliansky *et al.,* 2003). The functions of some PLRV proteins are known, but for others there are just speculations (Mayo and Ziegler-Graff, 1996). p28, a translation product of ORF0, has been suggested to play a role in host recognition and viral symptom expression (van der Wilk *et al*., 1997a). P1, a translation product of ORF1, has proteinase and helicase activities and contains a genome-linked protein, VPg (van der Wilk *et al*., 1997b). P2, a translation product

*Corresponding author. E-mail: hatem.fakhfakh@fsb.rnu.tn, fax: +21670-860432.

Abbreviations: BWYV = Beet western yellow virus; CYDV-RPV = Cereal yellow dwarf virus RPV; PLRV = Potato leafroll virus

of ORF2, is an RNA-dependent RNA polymerase. The viral capsid is composed of two proteins: a major coat protein encoded by ORF3 and a minor readthrough protein expressed from ORF3, which is believed to be an important aphid-transmission determinant of the virus (Chay *et al*., 1996; Ziegler-Graff *et al*., 1996). The putative proteins encoded by ORF6 and ORF7 were so far not identified except for P7, which displays nucleic acid-binding properties (Ashoub *et al*., 1998). P4 is a putative movement protein.

In contrast to some other poleroviruses, such as, isolates of PLRV differ in sequence only a little from each other. The identity of 12 PLRV isolates was 94–98% over all ORFs and comparisons among ORF3 sequences of these and a further seven isolates revealed no differences (Guyader and Ducray, 2002). Similarly, the Peruvian isolates varied only a little from each other or from others (Taliansky *et al*., 2003). The genomic organization of PLRV is very similar to those of the other poleroviruses, such as BWYV (Veidt *et al*., 1988). Comparisons made between the PLRV and BWYV proteins revealed that all they shared a high homology except for the ORF0 protein (p28). Although the ORF0s of both viruses are similar in size and position on the genome, their p28 sequences differ (Veidt *et al*., 1988).

Since PLRV is in Tunisia one of the main enemies for potato crop with about 35% losses (Djilani Khouadja *et al*., 2003), the methods capable of both detecting the virus in very low concentrations and discriminating between different isolates/strains should be available. PCR and RT-PCR techniques have become methods of choice for theirs specificity and sensitivity. Linking RT-PCR to automated DNA sequencing was shown to be very helpful for determination of the relatedness of isolates/strains at nucleotide level.

In this study, the variability of ORF0 of Tunisian isolates of PLRV originating from three different geographic regions and other isolates was assessed. Also a phylogenetic tree, based on the deduced amino acid sequences of ORF0 of various PLRV isolates was constructed. Finally, we searched for a relationship between the ORF0 sequence and the geographical origin or bioclimatic zone of the Tunisian isolates.

Materials and Methods

Virus isolates. A total of nine PLRV isolates originating from potato (*Solanum tuberosum* ssp. *tuberosum*) cv. Spunta from three geographical and bioclimatical areas, Mahdia (Central-East Tunisia, arid bioclimate), Bizerte (North Tunisia, upper semi-arid bioclimate) and Tunis (North-East Tunisia, subhumid bioclimate) were maintained on *Physalis floridana* plants (Djilani Khouadja *et al*., 2004) and were propagated by cuttings (Emberger, 1966) (Table 1). All isolates were collected in 2002.

Table 1. Characteristics of Tunisian PLRV isolates

Isolate	Place of origin	Bioclimate	Acc. No.
is_5	Mahdia	Arid	AY645677
is_6			AY645682
is_{SAM}	Bizerte	Sub-humid	AY645685
is_4	Mahdia	Arid	AY645681
is_{16}			AY645678
is_{24}	Tunis	Upper semi-arid	AY645683
is_{26}			AY645679
is_{29}			AY645684
is_{30}			AY645680

Table 2. The primers used in RT-PCR and sequencing

Primer	Assay	Sequence	Position (nts)
ORFI S2	RT-PCR	5'-GAAATTGCAGCTTTAG-3'	116–132
ORFI AS		3'-CATGTCACCTCTCTTGCGGA-5'	1020–1040
290S	Sequencing	5'-GCAAGCGAGCTTAATTTACG-3'	290–310
590S		5'-CGGTTGGGGACAAGACATG-3'	590–609
890S		5'-GGGCAGCTTTCTCCAAAAGG-3'	890–910
330AS		3'-CTAAATGGCGTATACGGCGAG-5'	309–330

Total RNA was extracted from infected *P. floridana* leaf samples using the RNeasy Plant Mini Kit (Qiagen) (Guyader and Giblot, 2002).

RT-PCR. The RT reaction mixture consisted of 13 µl (100 µg) of total RNA and 7 µl of a mixture containing the 1x buffer, 1 mmol/l dNTPs, 1.25 µmol/l ORF1 AS primer (Table 2), 20 U of RNasin, and 3 U of AMV reverse transcriptase (Promega). The reaction run at 42°C for 1 hr, yielding cDNA. In the PCR step, a 934 nt region encompassing the entire ORF0 was amplified using specific primers ORF1 S2 and ORF1 AS. The PCR reaction mixture consisted of 2.5 µl of cDNA and 22.5 µl of a mixture containing 0.5 µmol/l primers, 0.4 mmol/l dNTPs, 1 U of *Taq* DNA polymerase (Promega), 2 mmol/l $MgCl_2$, and the 1x free $MgCl_2$ PCR buffer. The reaction consisted of initial denaturation at 94°C for 4 mins, 35 cycles of 94°C/1 min, 57°C/1 min, and 72°C/1 min, and final extension final extension at 72°C for 10 mins in a thermocycler (Crocodile Oncor, Appligene).

Nucleotide sequencing of purified PCR products was carried out directly using the Concert Rapid PCR Purification System (Gibco-BRL) and specific primers 290S, 590S, 890S and 330AS (Table 2). The obtained sequences were analyzed using an ABI 310 automated sequencer (Applied Biosystems).

Sequence analysis. The sequences were assembled and aligned using the CAPCONTIG (Huang, 1992) and CLUSTAL W 1.8 (Thompson *et al.*, 1994) programs, respectively, both available in the BIOEDIT package (Hall, 1999). In determining the variability of individual nucleotide or amino acid positions in ORF0 within Tunisian isolates of PLRV, the Scotish PLRV-S isolate was used as reference (Mayo *et al.,* 1989). The identity of individual Tunisian isolates and that of Tunisian with other (GenBank) isolates were assessed pairwise using the maximum likelihood approach.

The tree was rooted by an outgroup (CYDV-RPV). The sequences of standard PLRV strains and isolates were retrieved from GenBank (Guyader and Giblot, 2000). Correlation between ORF0 sequences and geographical distance was looked for by the Mantel's test (Mantel, 1967) using the TFPGA 1.3 program (Miller, 1997). The significance of correlation was estimated at $P < 0.05$.

Results

ORF0 sequence analysis of Tunisian PLRV isolates

Based on the ORF0 nucleotide sequence, Tunisian PLRV isolates were compared with the Scottish PLRV-S isolate as reference. Among Tunisian isolates, 57 differences were dispersed all over the ORF0 region, 5 of which were common to all isolates; 3 of the 5 common were silent. The most variable nucleotide positions were concentrated in the region 305–414 and most (~76%) of the differences were silent. The is_5 and is_6 isolates showed identical ORF0 sequence.

Analysis of the deduced amino acid sequences showed differences in 15 positions, two of which (Tyr57 and P102) were common to all Tunisian isolates. At the nucleotide level, most common were transitions, namely C→U in 40%, which were also observed by Palucha *et al.* (1994) and de Souza-Dias *et al.* (1999). Most Tunisian isolates showed 42–45 nucleotide differences in comparison to PLRV-S, except for is_{SAM} (15) and is_5 and is_6 (13).

The identity of individual TU isolates, based on the ORF0 sequence, was also assessed. At the nucleotide level, they shared 92.42–99.73% identity (average 95.5%), regardless of their geographic origin. At the amino acid level, similarly, they shared 92.3–99.59% identity (average 96.4%). On the other hand, is_{26} and is_{SAM} appeared be the most divergent isolates with a 92.3% identity.

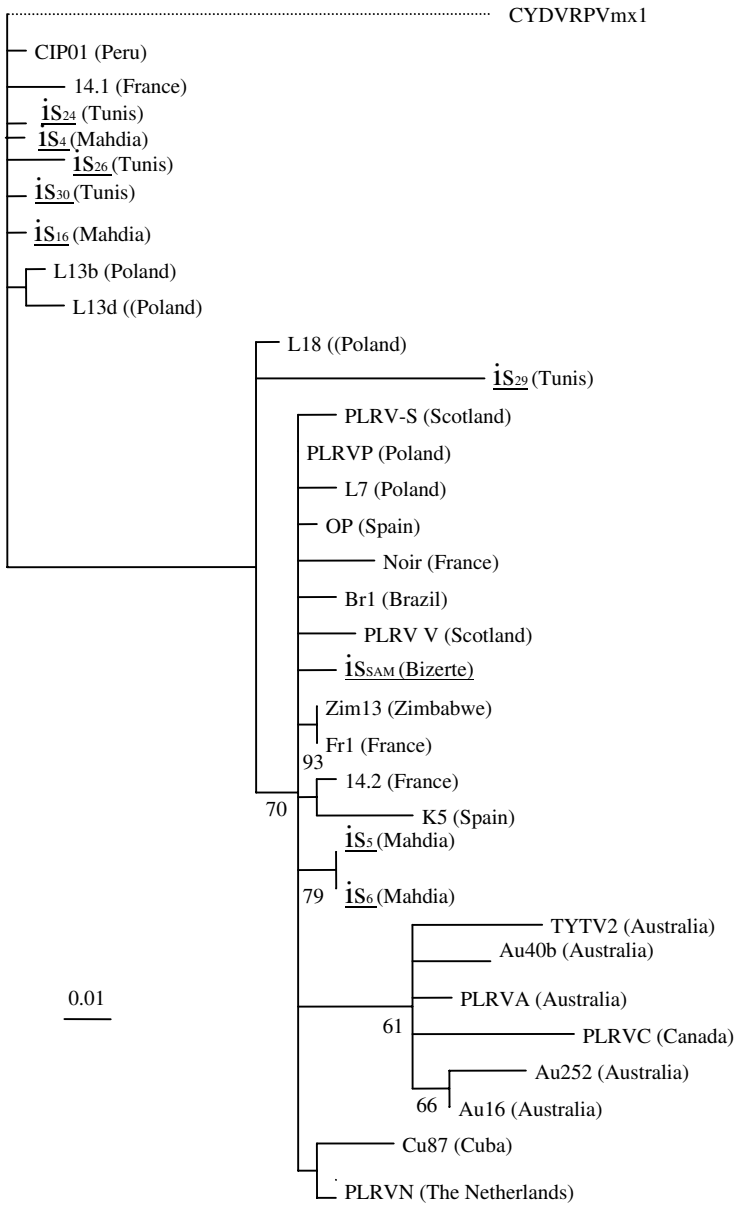

Fig. 1

Phylogenetic tree of PLRV isolates based on the amino acid sequence of ORF0 (P0) protein

Comparison of Tunisian with other PLRV isolates

Comparison of Tunisian with other PLRVs isolates showed 89.94–99.45% identity (average 95.5%) at nucleotide level and 89.8–99.59% identity (average 95.0%) at amino acid level. Tunisian isolates is_5, is_6 and is_{SAM} were most closely related to others with 96.8% identity. Furthermore, Tunisian isolates were most similar to Peruvian isolates (~97%) but less similar to the isolates from Poland (~95%), France, Netherlands, Scotland, Cuba, and Zimbabwe (all 94%), Spain and Brazil (both ~93%), Australian and Canada (~92%). Furthermore, this comparison revealed 5 conserved regions of 16–29 bp.

Phylogenetic analysis

The phylogenetic relationships between Tunisian and other PLRV isolates were assessed by constructing a phylogenetic tree based on the ORF0 amino acid sequence (Fig. 1). The tree showed two major clusters, the second consisting of three subclusters. Tunisian isolates did not cluster together but were distributed more widely, regardless of their geographic origin: four isolates were located within the first major cluster, while remaining five isolates occurred in the first subcluster of the seconf major cluster. Another comparative analysis did not prove any correlation between the amino acid sequence and the bioclimatic zone of the origin of Tunisian isolates.

Discussion

In this study, analysis of ORF0 sequences revealed a moderate level of similarity of Tunisian PLRV isolates to those from other parts of the world. On the other hand, Tunisian isolates collected from the same region were no less divergent from each other than were those collected from different regions. Since ORF0 was known to play a role in symptom expression (van der Wilk *et al.*, 1997), its sequence might correlate with the disease symptoms severity on *P. floridana,* as investigated in this work. In fact, the isolates is_5, is_6 and is_{SAM} clustered together and induced very severe symptoms on *P. floridana* test plants (data not shown). However, the possible correlation, in general, was not proved.

The use of PCR to detect different pathogens of plants has become a powerful diagnostic tool. Depending on the choice of primers, PCR can provide both narrow and broad specificities for various isolates or strains of pathogens (Hadidi *et al*., 1995; Henson and French, 1993). Here, five conserved regions among PLRV sequences were found. More consensus sequences were detected in other genome regions (Haliloglu and Bostan, 2002). It could be very informative to locate regions of several sequences, which are well conserved to design primers for detection of viruses or isolates by PCR.

In this work, the ORF0 sequence analysis revealed that possibility to distinguish some isolates from others in accord with Guyader and Giblot (2002). On the other hand, these authors showed that the genetic diversity of PLRV, as assessed by analyzing genomic sequences of isolates originating from different countries and continents, was low. All this confirms that the ORF0 products constitute the main genetic difference between PLRV isolates/strains and represents one of its most variable coding regions. In contrast, comparisons of PLRV isolates from around the world (Keese *et al*., 1990; Palucha *et al*., 1994) have shown that the isolates from Canada, Poland, Scotland and the Netherlands differ in nucleotide sequence over the whole genome by about 2%. An Australian isolate differed from all these isolates by about 7%. In a comparative study based on P3 and P5 genes, Jolly (1994) has found that Scottish isolates were as divergent (1%) from each other as from overseas isolates except that from Australia (Keese *et al*., 1990); summing up, this study showed a low variability of PLRV.

The phylogenetic tree, based on ORF0 protein, illustrated the relationship between Tunisian and other isolates of PLRV. Interestingly, PLRV isolates in general did not cluster according to their geographic origin. In particular, Tunisian isolates clustered in three different groups. These findings are basically in accord with those of Guyader and Ducray (2002), though the respective trees were not identical. The lack of sequence variation suggests either that PLRV only recently diverged from an ancestral virus, e.g. by acquiring an ability to infect potato, or it has been subject to very strong selection constraints. The narrow genetic base of potatoes in cultivation may also represent a factor in restricting PLRV variation (Guyader and Ducray, 2002).

Sequence variability may have important implications in use of the P0 gene for transgenic resistance. Several lines of evidence from different host systems suggest that such a resistance can be highly sequence-specific (Gonsalves, 1998; Nakajima *et al*., 1993; Nelson *et al*., 1988; Sanders *et al*., 1992). This study provides preliminary information about the variabilty of ORF0 in PLRV isolates for development of the strategy of designing as well as engineering the ORF0-based transgenic resistance in plants. In conclusion, unlike other PLRV genome regions, ORF0 was found to be not highly conserved among PLRV isolates both at nucleotide and amino acid level.

Acknowledgements. This work was partially supported by grants from the Ministry of High Education, the Ministry of Scientific Research, Technology and Competence Development, the CNRS of France, the French Embassy in Tunisia, and a grant from the Middle East Regional Cooperation. The authors thank Dr. S. Guyader for technical help and critical comments to the manuscript.

References

Ashoub A, Rohde W Prufer D (1998): In planta transcription of a second subgenomic RNA increases the complexity of the subgroup 2 luteovirus genome. *Nucleic Acids Res.* **26**, 420–426.

Chay CA, Gunasinge UB, Dinesh-Kumar SP, Miller WA, Gray SM (1996): Aphid transmission and systemic plant infection determinants of barley yellow dwarf luteovirus-PAV are contained in the coat protein readthrough domain and 17-kDa protein, respectively. *Virology* **219**, 57–65.

Djilani Khouadja F, Guyader S, Gorsane F, Khamassy N, Rouzé J, Marrakchi M, Fakhfakh H (2003): Diagnosis and

molecular analysis of Potato leafroll virus isolates in Tunisia. *EPPO Bull.* **33**, 361–368.

Djilani Khouadja F, Rouzé-Jouan J, Gauthier JP, Bouhachem S, Marrakchi M, Fakhfakh H (2004): Transmission efficiency of Tunisian Potato leaf-roll virus isolates by Tunisian clones of the *Myzus persicae* complex (Hemiptera: Aphididae). *Bol. San. Veg. Plag.* **30**, 47–55.

Emberger L (1966): Une classification biogéographique des climats. Recherches et Travaux des Laboratoires de Géologie, Botanique et Zoologie. Vol. 7, Faculté des Sciences Montpellier, France, pp. 1–43.

Gonsalves D (1998): Control of papaya ringspot virus in papaya: A case study. *Annu. Rev. Phytopathol.* **36**, 415–437.

Guyader S, Giblot-Ducray D (2002): Sequence analysis of potato leafroll virus isolates reveals genetic stability, major evolutionary events and differential selection pressure between overlapping reading frame products. *J. Gen. Virol.* **83**, 1799–1807.

Hadidi A, Levy LV, Podleckis EV (1995): Polymerase chain reaction technology in plant pathology. In Singh RP, Singh US (Eds): *Molecular Methods in Plant Pathology.* CRC/Lewis Press, Boca Raton, pp. 167–187.

Haliloglu K, Bostan H (2002): Nucleotide Sequence Analysis for Assessment of Variability of Potato Leafroll Virus and Phylogenetic Comparisons. *J. Biol. Sci.* **2**, 582–586.

Hall TA (1999): BioEdit: a user friendly biological sequence alignment editor and analysis program for Window 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98.

Harrison BD (1984): Potato leafroll virus. *Commonwealth Mycological Institute/Association of Applied Biologists* (*CMI/AAB*) *Descriptions of Plant Viruses*, p. 291.

Henson JM, French R (1993) The polymerase chain reaction and plant disease diagnosis. *Annu. Rev. Phytopathol.* **331**, 81–109.

Huang XA (1992): Contig assembly program based on sensitive detection of fragment overlaps. *Genomics* **14**, 18–25.

Keese P, Martin RR, Kawchuk LM, Waterhouse PM, Gerlach WL (1990): Nucleotide sequences of an Australian and a Canadian isolate of potato leafroll luteovirus and their relationships with two European isolates. *J. Gen. Virol.* **71**, 719–724.

Mantel N (1967): The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220.

Mayo MA, Robinson DJ, Jolly CA, Hyman L (1989): Nucleotide sequence of potato leafroll luteovirus RNA. *J. Gen. Virol.* **70**, 1037–1051.

Mayo MA, Ziegler-Graff V (1996): Molecular biology of luteoviruses. *Adv. Virus Res.* **46**, 416–460.

Miller MP (1997): Tools for population genetic analysis (TFPGA) 1.3: A Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by authors.

Nakajima M, Hayakawa T, Nakamura I, Suzuki M (1993): Protection against cucumber mosaic virus (CMV) strain O and Y and chrysanthemum mild mottle virus in transgenic tobacco plants expressing CMV-O coat protein. *J. Gen. Virol.* **74**, 319–322.

Nelson RS, McCormick SM, Delannay X, Dube P, Layton J, Anderson EJ, Kaniewska M, Proksch RK, Horsch RB, Rogers SG, Fraley RT, Beachy RN (1988): Virus tolerance, plant growth and field performance of transgenic tomato plants expressing coat protein from tobacco mosaic virus. *Biotechnology* **6**, 403–409.

Palucha A, Sadowy E, Kujawa A, Juszczuk M, Zagorski W, Hulanicka D (1994): Nucleotide sequence of RNA of a Polish isolate of potato leafroll luteovirus. *Acta Biochim. Pol.* **41,** 405–414.

Robert Y, Lemaire O (1999): Epidemiology and control strategies. In Smith HG, Barker H (Eds): *The Luteoviridae*. CABI Publishing, CAB International, Wallingford, UK, pp. 211–279.

Sanders PR, Sammons B, Kaniewski W, Haley L, LaVallee BJ, Delannay X, Turner NE (1992): Field resistance of transgenic tomatoes expressing the tobacco mosaic virus or tomato mosaic virus coat protein genes. *Phytopathology* **82**, 683–690.

Shepardson S, Esau K, McCrum R (1980): Ultrastructure of Potato leaf phloem infected with potato leafroll virus. *Virology* **105**, 379–392.

de Souza-Dias JAC, Russo P, Miller L, Slack A (1999): Comparison of nucleotide sequences from three potato leafroll (PLRV) isolates collected in Brazil. *Am. J. Potato Res.* **76**, 17–24.

Taliansky M, Mayo MA, Barker H (2003): Potato leafroll virus: a classic pathogen shows some new tricks. *Mol. Plant Pathol.* **4,** 81–89.

Thomas JE (1993): Alternative hosts and epidemiology of potato leafroll virus in Queensland. *Aust. J. Agric. Res.* **44**, 1905–1916.

Thompson JD, Higgins DG, Gibson TJ (1994): CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.

van der Wilk F, Houterman P, Molthoff J, Hans F, Dekker B, Van den Heuvel JFJM, Huttinga1 H, Goldbach R (1997a): Expression of the Potato Leafroll Virus ORF0 Induces Viral-Disease-like Symptoms in Transgenic Potato Plants. *Am. Phytopathol. Soc.* **10**, 153–159.

van der Wilk F, Verbeek M, Dullemans AM, Van den Heuvel JFJM (1997b): The genome-linked protein of potato leafroll virus is located downstream of the putative protease domain of the ORF1 product. *Virology* **234**, 300–303.

Veidt I, Lot H, Leiser M, Scheidecker D, Guilley H, Richards K, Jonard G (1988): Nucleotide sequence of beet western yellows RNA. *Nucleic Acids Res.* **16**, 9917–9932.

Waterhouse PM, Gildow FE, Johnstone GR (1987): Luteovirus group. In *Descriptions of Plant Viruses*. Common. Mycol. Inst./Assoc. Appl. Biol., Kew, England, No. 339.

Ziegler-Graff V, Brault V, Mutterer JD, Simonis MT, Herrbach E, Guilley H, Richards KE, Jonard G (1996): The coat protein of beet western yellows luteovirus is essential for systemic infection but the viral gene products P29 and P19 are dispensable for systemic infection and aphid transmission. *Mol. Plant-Microbe Interact.* **9**, 501–510.