Identification of a pathogenicity determinant of *Plum pox virus* in the sequence encoding the C-terminal region of protein P3+6K₁

Pilar Sáenz,1 Maria Teresa Cervera,1 Silvie Dallot,2 Laurence Quiot,2 Jean-Bernard Quiot,2 José Luis Riechmann†1 and Juan Antonio García1

1Centro Nacional de Biotecnología (CSIC), Campus de la Universidad Autónoma de Madrid, 28049 Madrid, Spain
2Ecole Nationale Superieure Agronomique de Montpellier (ENSAM-INRA), 2 Place Viala, 34060 Montpellier Cedex 1, France

A full-length genomic cDNA clone of a plum pox potyvirus (PPV) isolate belonging to the M strain (PPV-PS) has been cloned downstream from a bacteriophage T7 polymerase promoter and sequenced. Transcripts from the resulting plasmid, pGPPVPS, were infectious and, in herbaceous hosts, produced symptoms that differed from those of virus progeny of pGPPV, a full-length genomic cDNA clone of the D strain PPV-R. Viable PPV-R/-PS chimeric viruses were constructed by recombination of the cDNA clones *in vitro*. Analysis of plants infected with the different chimeras indicated that sequences encoding the most variable regions of the potyvirus genome, the P1 and capsid protein coding sequences, were not responsible for symptom differences between the two PPV isolates in herbaceous hosts. On the contrary, complex symptomatology determinants seem to be located in the central region of the PPV genome. The results indicate that a genomic fragment that encodes 173 aa from the C-terminal part of the P3+6K₁ coding region is enough to confer, on a PPV-R background, a PS phenotype in *Nicotiana clevelandii*. This pathogenicity determinant also participates in symptom induction in *Pisum sativum*, although the region defining the PS phenotype in this host is probably restricted to 74 aa.

**Introduction**

*Plum pox virus* (PPV) is a member of the large and economically important genus *Potyvirus* from the family *Potyviridae* (Simón et al., 1997). The potyvirus genome consists of a single-stranded messenger-polarity RNA molecule of about 10 kb, with a VPg protein at its 5′ end and a poly(A) tail at its 3′ end. It is translated into a large polyprotein that is processed proteolytically by three virus-encoded proteases (Riechmann et al., 1992; Revers et al., 1999).

In nature, PPV causes a very damaging disease of trees of the genus *Prunus*. Many PPV isolates have been classified in several groups according to serological and biological properties, particularly the symptoms caused in experimental herbaceous hosts (Kerlan & Dunez, 1979; Sutic et al., 1971;Van Oosten, 1971). The availability of genome sequence data has enabled a more reliable classification to be established. Apart from the atypical El Amar isolate (Wetzel et al., 1991) and the recently characterized isolates that infect sweet and sour cherries (Crescenzi et al., 1997; Nemchinov & Hadidi, 1996), most PPV isolates can be classified into two major strains, M (from the isolate Marcus) and D (from the isolate Dideron) (Candresse et al., 1998).

Full-length cDNA clones, from which infectious transcripts can be produced either *in vitro* or *in vivo* allowing easy manipulation of the viral genome, are very useful tools for the identification of strain- and isolate-specific pathogenicity determinants. Whereas biologically active full-length cDNA clones from the D isolates PPV-R (Riechmann et al., 1990) and PPV-NAT (Maiss et al., 1992) have been reported, full-length cDNA clones of PPV isolates belonging to the M strain have not been described. In this paper, we report the construction of...
a full-length cDNA clone of an M isolate of PPV, PPV-PS, and demonstrate the viability of M/D recombinant viruses. We have taken advantage of the different symptomatology caused in several herbaceous hosts by the virus progeny of the PPV clones to gain insight into the elements of the potyvirus genome that influence symptom development.

Methods

- **Virus and bacterial strains.** The PPV-PS isolate was obtained from M. Rankovic (Fruit Research Institute, Cacak, Yugoslavia) and maintained in *Nicotiana clevelandii* plants. Virus purification and RNA extraction were performed as described previously (Lain et al., 1988). *Escherichia coli* DH5α and JM109 were used for cloning of the plasmids and CJ236 was used in site-directed mutagenesis experiments.

- **Synthesis and cloning of cDNA from PPV-PS RNA.** The construction of clones of the p1 and pBS series, which contain partial cDNA copies of PPV-PS RNA inserted in the Hind III site (pH) or between the BamHI and SacI sites (pBS) of pUC18, has been reported previously (Cervera et al., 1993). Additional clones with fragments covering the entire PPV-PS genome except its 5' end were obtained in a similar fashion. In order to clone the first 547 nt of the PPV-PS genome, a cDNA fragment was amplified by RT-PCR using the oligodeoxynucleotides 5' AAAATATAAAAACTCAACAC 3' and 5' CCAGTGTGTGTACCG 3' as primers. This fragment was digested with HindIII or AclI and cloned into the pGEM3 vector (Promega) digested with HindIII and SmaI or AclI and SmaI, generating the plasmids pG5 'PSH and pG5 'PSA, respectively. Several independent clones were sequenced to ensure that no mutation had been introduced during the RT-PCR. Non-viral nucleotides between the transcription initiation site of the pGEM3 T7 promoter and the first nucleotide of PPV were removed by site-directed mutagenesis (Kunkel et al., 1987) by using the oligodeoxynucleotide 5' TTTATATTTTCTAGTGAC 3' (Riechmann et al., 1990), to give pG5 'PSH.

The complete nucleotide sequence of the PPV-PS genome was determined by sequencing double-stranded cDNA contained in the different plasmids and single-stranded DNA subcloned into M13 recombinant phages with the Sequenase 2.0 kit (Amersham). The first 211 nt of the PPV-PS genome were determined by direct sequencing of viral RNA (Fichot & Girard, 1990).

The complete clone pGPPVPS was prepared by replacing appropriate restriction fragments in pGPPV, the previously described full-length cDNA of PPV-R (Rankovic) (Riechmann et al., 1990), with corresponding fragments from PPV-PS. The PPV-PS/PPV-R chimeras shown in Fig. 1 were produced by exchange of restriction fragments as indicated in the figure. All fragments generated by PCR were sequenced. Additional details will be supplied by the authors on request.

- **In vitro transcription and inoculation of plants.** The recombinant plasmids were linearized prior to transcription by digestion either with *Pvu*II and *Bst*I or with *Pvu*II alone, followed by phenol–chloroform extraction and ethanol precipitation. Capped full-length transcripts were synthesized from these templates by using the T7 CapScribe kit (Boehringer Mannheim). Yield and integrity of the transcripts were analysed by agarose gel electrophoresis.

Three primary leaves per *N. clevelandii* plant were dusted with carborundum and inoculated mechanically with 1-3 μl of the transcription reaction mixture diluted 1:1 with 5 mM sodium phosphate buffer (pH 7.2). Crude sap from leaves of infected *N. clevelandii* plants (2 ml per g tissue) was used to inoculate *Psium sativum* plants and *N. clevelandii* plants (20 μl in three leaves per plant) by hand.

Virus accumulation was assessed by Western blot and by double-antibody sandwich indirect (DASI)-ELISA with the REALISA kit (Durviz). Fragments that included the PS–R borders were amplified by RT–PCR (Titan kit; Boehringer Mannheim) preceded by immunocapture (Wetzel et al., 1992) and sequenced to verify that the chimeric sequences were maintained in virus progeny.

Results

**Virus progeny of infectious transcripts synthesized from a PPV-PS full-length cDNA copy cause mild symptoms in *N. clevelandii***

cDNA cloning and partial sequencing of the PPV-PS isolate have been reported previously (Cervera et al., 1993). We have now determined the complete genomic sequence of 9786 nt [excluding the 3’-terminal poly(A) sequence] of PPV-PS. PPV-PS has very low nucleotide (12.2%) and amino acid (19.9%) divergence from the other M-type isolate that has been sequenced completely, PPV-SK 68 (Palkovics et al., 1993). Its level of divergence from sequenced D-type isolates (Lain et al., 1989; Maiss et al., 1989, 1991; Teycheney et al., 1989) is much higher, ranging from 12.5 to 12.9% in nucleotide sequence and from 4.2 to 4.8% in amino acid sequence. There was little difference in the degree of sequence variation along the genome, except for the sequence encoding the first 270 nt of the capsid protein (CP) gene, where the PPV-PS sequence diverged from the sequence of PPV-SK 68 and the D-type isolates by 1.9 and 28.1–28.9%, respectively, corresponding to levels of amino acid divergence of 2.1% (PPV-PS vs -SK 68) and 31.9–32.4% (PPV-PS vs D-type isolates).

A full-length cDNA copy of the PPV-PS isolate was cloned downstream from a phage T7 RNA polymerase promoter. Capped RNA transcripts, synthesized *in vitro* from the resulting plasmid pGPPVPS, displayed a level of infectivity in *N. clevelandii* plants similar to that observed for transcripts of pGPPV, a full-length cDNA clone of the Rankovic isolate of PPV (PPV-R). Infection with the virus progeny of PPV-PS transcripts (PPV-PSc) caused a very mild chlorotic mottle, clearly distinguishable from the more severe symptoms provoked by virus progeny of PPV-R transcripts (PPV-Rc) (Fig. 2A). The mild symptoms of PPV-PSc-infected plants resembled those caused by several virus variants isolated from the original PPV-PS isolate, but largely differed from those of other variants, indicating the existence of several sub-isolates with different biological properties (P. Sáenz & J. A. García, unpublished results).

A genome fragment that includes the P3+6K, cistron and part of the HC and CI cistrons is sufficient to confer R-type symptomatology in *N. clevelandii* plants

In order to map PPV genomic sequences involved in symptom determination, chimeric recombinants were constructed between the parental full-length cDNA clones,
pGPPVPS and pGPPV. The first chimeras focused on the CP and P1 coding regions, the most variable regions among potyviruses (Aleman-Verdaguer et al., 1997). In addition, molecular features of the CP cistron are the current discrimination criteria used to classify PPV isolates (Bousalem et al., 1994). Fig. 3 shows a schematic representation of the hybrid clones and summarizes the symptomatology produced by them. Virus progeny derived from pR/P 7677 produced R-type symptomatology, although the complete CP cistron of PPV-PS was present. Moreover, virus derived from pR/P 2212 and pR/P 2212–7677, which include P1 and P1 plus CP coding sequences, respectively, from PPV-R in the PPV-PS back-
Fig. 2. (A) Symptoms observed in detached leaves of *N. clevelandii* infected with virus progeny of pGPPV (type R symptoms, at
Potyvirus pathogenicity determinant

Virus accumulation in plants infected with the different chimeras was assessed by Western blot (Fig. 3) and ELISA (Fig. 4A). Although there were small differences in the levels of accumulation of the different chimeric viruses, no correlation was observed between symptom severity and virus accumulation.

Exchange of sequences encoding part of the C-terminal region of the P3+6K, protein converts the R-type to the PS-type phenotype

To map the smallest region of the PPV genome able to elicit PS-type symptoms, five additional chimeras in the 2212–3628 region were constructed (Fig. 5). The PS 2904–3628 fragment (pR/P 2904–3628), but not the 2212–2904 fragment (pR/P 2212–2904), conferred PS symptomatology to PPV-R (Fig. 2B). However, the R phenotype was maintained in the two chimeric viruses that contained, in the PPV-R background, the PS 2904–3409 (pR/P 2904–3409) and 3411–3628 (pR/P 3411–3628) complementary fragments, where the 2904–3628 region was divided (Fig. 2B). Increasing the size of the PS region of pR/P 3411–3628, in such a way that it started at nt 3109 (pR/P 3109–3628), led to the reappearance of PS symptomatology (Fig. 2B). No correlation
**Fig. 4.** Virus accumulation, estimated by DASI-ELISA, in extracts of *N. clevelandii* (A, B) and *P. sativum* (C) plants infected with virus progeny of pGPPV, pGPPVPS and the hybrid viruses described in Fig. 1. Constructs that led to PS-type symptoms are indicated by filled bars; those that led to R-type symptoms are indicated by open bars.

**Fig. 5.** Schematic representation of different chimeric viruses showing the type of symptoms that they produce in *N. clevelandii* and *P. sativum* plants. An immunoblot analysis of protein extracts from *N. clevelandii* plants infected with the different chimeric viruses is shown at the bottom of the figure. Samples (3 mg tissue from each infected plant) were collected at 21 days post-inoculation and used in the assay. The immunoreaction was carried out with an antiserum raised against the coat protein (CP) of PPV-R.
between virus accumulation, estimated by Western blot (Fig. 5) and ELISA (Fig. 4B), and symptom severity was observed in plants infected with these chimeras. The 3109–3628 region of the PPV genome, which contains all the information required to transform the R-type into the PS-type symptomatology, encodes 173 aa that differ at 11 positions between PPV-R and PPV-PS (Fig. 6).

The determinants governing the differences between PPV-R and PPV-PS symptomatology are similar but not identical in *N. clevelandii* and *P. sativum*

In *P. sativum*, PPV-R produced local necrotic lesions and systemic chlorotic spots that became necrotic. In contrast, PPV-PS only caused a systemic chlorotic mottling that did not become necrotic. As in the *N. clevelandii* infection, the sequences between nt 2212 and 7677 carried all the information required to define both the R- and PS-type symptomatologies (Fig. 3; symptoms not shown). However, in contrast to the situation in *N. clevelandii*, in *P. sativum*, not only the 2212–5535 fragment, but also a 2212–3628 fragment, was able to confer R-type symptomatology in a PS background (Fig. 3; symptoms not shown). This indicates that the determinants of R-type symptomatology located between positions 3628 and 5535 are important in one host (*N. clevelandii*), but not in the other (*P. sativum*). Moreover, virus derived from pR/P 3411–3628, which showed an R-type phenotype in *N. clevelandii*, caused PS-type symptoms in *P. sativum* (Fig. 5). Thus, in this host, it was possible to limit the genomic region required for symptom exchange to that encoding a 74 aa sequence of the P3 + 6K1 C-terminal region, which differs by only 4 aa between PPV-PS and PPV-R (Fig. 6). As in *N. clevelandii* plants, no correlation was observed between virus accumulation and symptom severity in *P. sativum* plants infected with the different chimeric viruses (Fig. 4C).

**Discussion**

Full-length cDNAs from which infectious transcripts can be produced *in vitro* (Riechmann *et al*., 1990) or *in vivo* (Maiss *et al*., 1992) have been described for two PPV isolates, PPV-R and PPV-NAT, respectively. These two isolates belong to the D strain. In this paper, we report the construction of a full-length cDNA from the PPV-PS isolate, which belongs to the second major PPV group, the M strain. Transcripts synthesized *in vitro* from this clone (pGPPVPS) were infectious and, in several herbaceous hosts, their virus progeny produced clearly distinct symptoms from those produced by virus progeny of transcripts synthesized from the full-length cDNA clone of PPV-R (pGPPV). We investigated the genetic determinants of these symptomatic differences by constructing chimeric R/PS constructs. In spite of the quite large nucleotide divergence between the PPV-PS and PPV-R isolates (12.5%), all chimeric viruses that were constructed were viable, demonstrating that genetic elements of both strains are compatible and that even hybrid proteins can be functional. It is important to mention that, although the results demonstrate that our approach can be useful to study strain-specific genetic determinants, the pathogenicity determinants that we analysed are isolate-specific rather than strain-specific, since different M-type isolates cause very different symptoms in herbaceous hosts.

Our results indicate that a genome fragment from the end of the P3 + 6K1 coding region is sufficient to confer a PS phenotype on a PPV-R background. The size of this pathogenicity determinant and the kind of symptoms that it induces depend on the host. In *N. clevelandii*, between 74 and 173 amino acids (including a maximum of 11 differences between PPV-PS and -R) near the C-terminal end of the PPV-PS P3 + 6K1 protein are necessary to get the mild systemic chlorotic mottle typical of pGPPVPS virus progeny. This same region from pGPPV does not define R-type symptoms in *N.
when cloned in pGPPVPS. Thus, the fact that virus progeny from pR/P 2212–3628 and pR/P 2212–5535 cause PS and R symptoms, respectively, indicates that specific CI sequences are required to elicit the severe symptomatology of the PPV-R isolate in *N. clevelandii*. It is important to note that, although very unlikely, effects due to differences at the nucleotide level cannot be ruled out.

The differences in symptomatology between PPV-R and -PS in *P. sativum* appear to be defined by the same pathogenicity determinant. However, in this case, a sequence encoding 74 amino acids of the C-terminal region of P3 + 6K<sub>1</sub> (containing four differences between PPV-PS and -R) is enough to elicit a PS-type response (systemic chlorotic spots without necrotic local lesions). Moreover, in this host, the pathogenicity determinants in the CI cistron are less important, since the 2212–3628 fragment from PPV-R, which does not include CI sequences, is enough to confer the typical phenotype of this isolate (local necrotic lesions and systemic chlorotic spots that become necrotic). It is important to remark that other genomic regions might define differences of pathogenicity between PPV-R and PPV-PS in other herbaceous hosts. In this regard, the necrotic (PPV-R)/chlorotic (PPV-PS) type of the local lesions that PPV causes in *Chenopodium foetidum* seems to be defined by complex determinants different from that described in this report. Thus, although R/P 2212–7677 (which has the central region of the genome of PPV-PS) caused chlorotic lesions in *C. foetidum*, the reciprocal clone, P/R 2212–7677, as well as most of the rest of chimeras, caused a confused pattern of chlorotic, necrotic and intermediate lesions (data not shown).

The fact that we did not observe any correlation between symptom severity and virus accumulation seems to indicate that the differences in symptomatology are not the simple result of different replication efficiencies of PPV-R and -PS. Several regions of the potyvirus genome have been shown to be involved in symptom determination: for example, the HC coding sequence of *Tobacco vein mottling virus* (TVMV) (Atreya et al., 1992), a genomic fragment that encodes the NIa protease and Nlb replicase (Johansen et al., 1996), the PPV 5' non-coding region (NCR) (Simón-Buela et al., 1997) and the 3' NCR of TVMV (Rodríguez-Cerezo et al., 1991). Interestingly, wilting symptoms in Tabasco pepper plants infected with *Tobacco etch virus* (TEV) are determined by two separate regions of the viral genome, one encoding the C-terminal part of the P3 + 6K<sub>1</sub> protein and the other encoding the C-terminal region of the CI protein, the 6K<sub>2</sub> peptide and the N-terminal region of the VPg protein (Chu et al., 1997). Thus, symptoms in potyvirus-infected plants seem to be the result of a complex interaction between different virus factors with the host, rather than the effect of a single virus protein. In agreement with this assumption, differences in symptomatology between PPV-R and -PS are determined by a large genomic fragment. Interestingly, the most variable potyviral proteins, P1 and CP, do not seem to play a role, whereas, as for the wilting phenotype of the TEV-infected Tabasco pepper, the C-terminal region of the P3 + 6K<sub>1</sub> protein is fundamental to the specific PPV phenotype in *N. clevelandii* and pea.

Little is known about the function of the P3 + 6K<sub>1</sub> protein in potyvirus infection. It is required for virus replication in the infected cell (Klein et al., 1994) and has been shown to interact with cytoplasmic (Rodríguez-Cerezo et al., 1993) and nuclear (Langenberg & Zhang, 1997) inclusions in TVMV and pea seed-borne mosaic virus-infected cells, respectively. The PPV P3 + 6K<sub>1</sub> protein is processed *in vitro* by the Nla protease (Garcia et al., 1992). Although processing at this site does not appear to be essential for virus viability, a mutation that disturbs the cleavage affects virus competitiveness, and the mutant form is either rapidly lost or the mutation is compensated for by a second mutation in the 6K<sub>1</sub> region that does not restore susceptibility to *in vitro* cleavage (Riechmann et al., 1995). From these data, it has been suggested that P3 + 6K<sub>1</sub> is probably the functional product and that detachment of the 6K<sub>1</sub> portion could have a regulatory role. Supporting our conclusion that P3 + 6K<sub>1</sub> is involved in symptom elicitation, it has been reported that stable mutations at the P3/6K<sub>1</sub> cleavage site that have little effect on *in vitro* cleavage, and that apparently do not affect virus accumulation, cause drastic changes in the symptoms of the infected plants, either alleviating or exacerbating them (Riechmann et al., 1995).

There is little information on how the P3 + 6K<sub>1</sub> protein may interact with plant and/or virus elements in order to elicit different symptomatologies. The TVMV P3 protein has been shown to be present predominantly in membrane-enriched fractions of extracts of infected leaves (Rodríguez-Cerezo & Shaw, 1991). Computer analysis indicates the existence of two putative transmembrane helices in the P3 domain. In addition, the 6K<sub>1</sub> peptide possesses a highly hydrophobic core and in this respect is very similar to the 6K<sub>2</sub> peptide, which has been shown to direct the Nla protein to membranes (Restrepo-Hartwig & Carrington, 1994). Thus, the P3 + 6K<sub>1</sub> protein may be an integral membrane protein and proteolytic removal of the hydrophobic 6K<sub>1</sub> domain may modulate its function. It is tempting to speculate that disturbance of cell membranes by P3 + 6K<sub>1</sub> insertion may contribute to symptom induction. On the other hand, the 6K<sub>1</sub> hydrophobic domain is flanked by hydrophilic sequences, and a mutation that removes a positive charge upstream of the hydrophobic region has been shown to be compensated for *in vivo* by downstream second mutations that recover the positive charge (J. L. Riechmann & J. A. García, unpublished data). This result suggests the involvement of electrostatic bonds in interactions between the C terminus of P3 + 6K<sub>1</sub> and host or virus factors, which could be modulated by 6K<sub>1</sub> cleavage and could be relevant to the development of symptoms in infected plants.

We wish to thank Elvira Domínguez for technical assistance and Juan José López Moya, Mª Rosario Fernández-Fernández and Carlos Málpica for critical reading of the manuscript. This work was supported by grants...
References


565


Received 16 August 1999; Accepted 8 November 1999