

Virus variants with differences in the P1 protein coexist in a *Plum pox virus* population and display particular host-dependent pathogenicity features

VARVARA I. MALIOGKA†,**, BEATRIZ SALVADOR‡,**, ALBERTO CARBONELL§, PILAR SÁENZ¶, DAVID SAN LEÓN, JUAN CARLOS OLIVEROS, MA OTILIA DELGADILLO, JUAN ANTONIO GARCÍA* AND CARMEN SIMÓN-MATEO*

Department of Plant Molecular Genetics, Centro Nacional de Biotecnología (CNB-CSIC), Campus Universidad Autónoma de Madrid, 28049 Madrid, Spain

SUMMARY

Subisolates segregated from an M-type *Plum pox virus* (PPV) isolate, PPV-PS, differ widely in pathogenicity despite their high degree of sequence similarity. A single amino acid substitution, K109E, in the helper component proteinase (HCPro) protein of PPV caused a significant enhancement of symptom severity in herbaceous hosts, and notably modified virus infectivity in peach seedlings. The presence of this substitution in certain subisolates that induced mild symptoms in herbaceous hosts and did not infect peach seedlings suggested the existence of uncharacterized attenuating factors in these subisolates. In this study, we show that two amino acid changes in the P1 protein are specifically associated with the mild pathogenicity exhibited by some PS subisolates. Site-directed mutagenesis studies demonstrated that both substitutions, W29R and V139E, but especially W29R, resulted in lower levels of virus accumulation and symptom severity in a woody host, *Prunus persica*. Furthermore, when W29R and V139E mutations were expressed concomitantly, PPV infectivity was completely abolished in this host. In contrast, the V139E substitution, but not W29R, was found to be responsible for symptom attenuation in herbaceous hosts. Deep sequencing analysis demonstrated that the W29R and V139E heterogeneities already existed in the original PPV-PS isolate before its segregation in different subisolates by local lesion cloning. These results highlight the potential complexity of potyviral populations and the relevance of the P1 protein of potyviruses in pathogenesis and viral adaptation to the host.

INTRODUCTION

Plum pox virus (PPV) is a potyvirus that causes sharka disease in stone fruit trees of the genus *Prunus* (García and Cambra, 2007). The PPV genome consists of a single-stranded RNA molecule, of 9786 nucleotides in length, which encodes a single polyprotein and a truncated frameshift product, which are proteolytically processed by three self-encoded proteases (Chung *et al.*, 2008; Salvador *et al.*, 2006). The replication of plus-stranded RNA genomes is carried out by RNA-dependent RNA polymerases that usually lack proofreading activity and thereby cause high rates of mutation (Malpica *et al.*, 2002). This, in combination with short replication times and high RNA yields, gives rise to complex and dynamic mutant clouds. This population structure is usually interpreted in terms of the quasispecies theory (Biebricher and Eigen, 2006; Domingo and Holland, 1997; Eigen, 1996), although other conceptual models have also been proposed (Holmes and Moya, 2002; Jenkins *et al.*, 2001). The mutant spectrum usually fluctuates around a unique consensus sequence. However, changes in selective pressure can cause quasispecies to evolve rapidly towards a different consensus sequence. Moreover, several dynamic virus populations, each centred on a consensus sequence, may coexist in single individuals. In these cases, the pathogenicity displayed is the result of complex intra- and interpopulation interactions and not necessarily a simple combination of the pathologies exhibited by the separate virus populations (Domingo and Holland, 1997). A complex population is probably an important factor in the success of viruses, because it allows adaptation to different hosts. Currently, little is known about the genetic basis for viral–host adaptation, although a few relevant analyses of potyvirus host range limitations have been reported (Ayme *et al.*, 2007; Charron *et al.*, 2008; Hajimorad *et al.*, 2011; Ohshima *et al.*, 2010; Salvador *et al.*, 2008a; Suehiro *et al.*, 2004).

The PPV-PS isolate belonging to the M strain was originally obtained from a severely infected peach tree in the former Yugoslavia, and propagated in *Nicotiana glauca*. Studies of local lesion cloning in *Chenopodium foetidum* revealed the complexity of the original virus population (Sáenz *et al.*, 2001). Several PPV-PS subisolates were segregated and found to share a high degree of sequence similarity (~99.9%) despite the manifestation

*Correspondence: Email: jagarcia@cnb.csic.es; csimon@cnb.csic.es

Present addresses:

†Plant Pathology Laboratory, School of Agriculture, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece;

‡Blood Systems Research Institute, 270 Masonic Ave., San Francisco, CA 94118, USA;

§Donald Danforth Plant Science Center, 975 N. Warson Rd., St Louis, MO 63132, USA;

¶PROGENIKA BIOPHARMA S.A., Edificio 801, Parque Tecnológico de Zamudio, 48160 Derio, Spain.

**These authors contributed equally to this work.

of large differences in pathogenicity in herbaceous and woody hosts (Sáenz *et al.*, 2001). Sequence analysis of the subisolates and site-directed mutagenesis of an infectious PPV-PS cDNA clone demonstrated that a single amino acid change (K109E) in the helper component proteinase (HCPro) protein produced a drastic enhancement of symptom severity in both *N. clelandii* and *Nicotiana occidentalis*. However, two mild subisolates, PPV-PS 1.3.1 and 2.1.1, resembled the severe subisolates in maintaining E109 in their HCPro sequence, suggesting that attenuating change(s) could exist in genomic regions of PPV-PS mild subisolates not yet sequenced (Sáenz *et al.*, 2001). We have now completed the genome sequencing of several PPV-PS subisolates and have identified two amino acid changes in the P1 protein involved in symptom attenuation in herbaceous hosts and loss of infectivity in peach seedlings.

RESULTS

PPV-PS 1.3.1 and 2.1.1 mild subisolates have two amino acid changes in their P1 protein

The complete genomic sequences of PPV-PS 1.3.1 and 2.1.1, two mild subisolates segregated from the original PPV-PS isolate by three local lesion passages in *C. foetidum* (Sáenz *et al.*, 2001), were determined on partial cDNA fragments amplified by reverse transcription-polymerase chain reaction preceded by immunocapture (IC-RT-PCR) from infected tissue. Only two nucleotide changes, which did not affect the amino acid sequence [G40A in the 5' noncoding region and G4250A in the cylindrical inclusion (CI) coding region], were found between these two subisolates (Table 1). Apart from these two nucleotides, the sequences of PPV-PS 1.3.1 and 2.1.1 differed from the revised cDNA sequence of the previously reported pGPPV-PS clone (PS-MCI) (Sáenz *et al.*, 2000) in six nucleotide positions (Table 1). In addition to the previously described K109E and S232G substitutions in HCPro,

PPV-PS 1.3.1 and 2.1.1 showed two nonconservative amino acid substitutions in the P1 protein (W29R and V139E) and one conservative amino acid substitution in the genome-linked viral protein (VPg) (S40T), with respect to the PS-MCI sequence. These substitutions were not shared by three PPV-PS subisolates that caused severe symptoms in *Nicotiana* plants and were able to infect peach seedlings: 4.1.4, which was sequenced some years ago and now has been partially resequenced, and 5.1.3 and 10₇, whose genomic sequences have been determined in this work (Table 1).

The amino acid change V139E affects symptom development in *N. clelandii* and *N. occidentalis*

In order to assess the significance of the virus variants with different P1 proteins found in the PPV-PS population, the amino acid substitutions were incorporated, independently or in combination, into PPV-PS cDNA constructs. As the viral progeny of the original PS-MCI clone only produced mild symptoms in herbaceous hosts, as a consequence of K109 in the HCPro protein (Sáenz *et al.*, 2001), this clone was not suitable for studying the ability of variants to attenuate viral symptoms. Instead, a PPV-PS cDNA clone, named pICPPV-PSeg (PSeg), was constructed containing the coding sequence of HCPro with E109. The expression of the viral sequence of PSeg is under the control of a cauliflower mosaic virus (CaMV) 35S promoter, which allows the production *in planta* of infectious transcripts after inoculation with plasmid DNA (López-Moya and García, 2000). Inoculation with PSeg caused severe symptoms in *N. clelandii* and *N. occidentalis* that were indistinguishable from those induced by the severe PPV-PS subisolates (Figs 1 and 2; Sáenz *et al.*, 2001). PSeg was further modified by site-directed mutagenesis to express P1 protein variants individually, i.e. PSegW29R and PSegV139E. These two clones were tested for their ability to infect *N. clelandii* plants (Fig. 1). Both clones infected 100% of biologically inoculated plants in two independent experiments. Although inoculation with PSegW29R caused

Table 1 Genome localization of nucleotide and amino acid differences between different *Plum pox virus* (PPV)-PS subisolates and the cDNA clone PSeg.

	5'NCR	P1				HCPro				CI	VPg			Nla-Pro		Nlb		CP				
Virus*	40	231 (29)	562 (139)	567 (141)	677 (177)	1395 (109)	1489 (140)	1764 (232)	1928 (286)	3654 (2)	4250 (200)	5832 (40)	6200 (162)	6854 (187)	6899 (202)	7212 (64)	7793 (257)	8711 (45)	8783 (69)	8840 (88)	9107 (177)	9242 (222)
<i>PS MCI</i>	G	T (W)	T (V)	T (F)	G (R)	A (K)	T (L)	A (S)	A (E)	T (L)	G (L)	T (S)	G (E)	T (D)	A (Q)	T (L)	A (S)	C (F)	G (P)	G (L)	C (G)	T (I)
PSeg†	G	T (W)	T (V)	T (F)	G (R)	G (E)	T (L)	G (G)	A (E)	T (L)	G (L)	T (S)	G (E)	T (D)	A (Q)	T (L)	A (S)	C (F)	G (P)	G (L)	C (G)	T (I)
4.1.4.	G	T (W)	T (V)	C (L)	A (R)	G (E)	T (L)	A (S)	G (E)	G (V)	G (L)	T (S)	G (E)	G (E)	A (Q)	T (L)	C (S)	A (L)	G (P)	G (L)	C (G)	A (I)
5.1.3.	G	T (W)	T (V)	T (F)	A (R)	G (E)	T (L)	A (S)	A (E)	G (V)	G (L)	T (S)	G (E)	G (E)	A (Q)	T (L)	C (S)	C (F)	G (P)	G (L)	C (G)	A (I)
<i>1.3.1.</i>	A	C (R)	A (E)	T (F)	G (R)	G (E)	T (L)	G (G)	A (E)	T (L)	A (L)	A (T)	A (E)	T (D)	A (Q)	T (L)	A (S)	C (F)	G (P)	G (L)	C (G)	T (I)
<i>2.1.1.</i>	G	C (R)	A (E)	T (F)	G (R)	G (E)	T (L)	G (G)	A (E)	T (L)	G (L)	A (T)	A (E)	T (D)	A (Q)	T (L)	A (S)	C (F)	G (P)	G (L)	C (G)	T (I)
10₇	G	T (W)	T (V)	T (F)	A (R)	G (E)	C (S)	A (S)	A (E)	G (V)	G (L)	T (S)	G (E)	G (E)	G (Q)	C (L)	C (S)	C (F)	A (P)	A (L)	T (G)	A (I)

CI, cylindrical inclusion protein; CP, coat protein; HCPro, helper component proteinase; NCR, noncoding region; Nla-Pro, nuclear inclusion protein 'a' proteinase; Nlb, nuclear inclusion protein 'b'; P1, protein 1; VPg, genome-linked viral protein.

Bold and italic rows indicate virus causing severe and mild symptoms, respectively.

*Top row indicates nucleotide position in the genome and, in parentheses, the amino acid position in each protein.

†PSeg also differed from all subisolates in three nucleotide changes at positions 2921, 2927 and 2936 which do not alter the amino acid sequence, and were introduced during the construction of the cDNA clone to facilitate the cloning of an intron (López-Moya and García, 2000), and in an extra T in the 3'NCR, at position 9629, which was introduced accidentally during the cloning procedure.

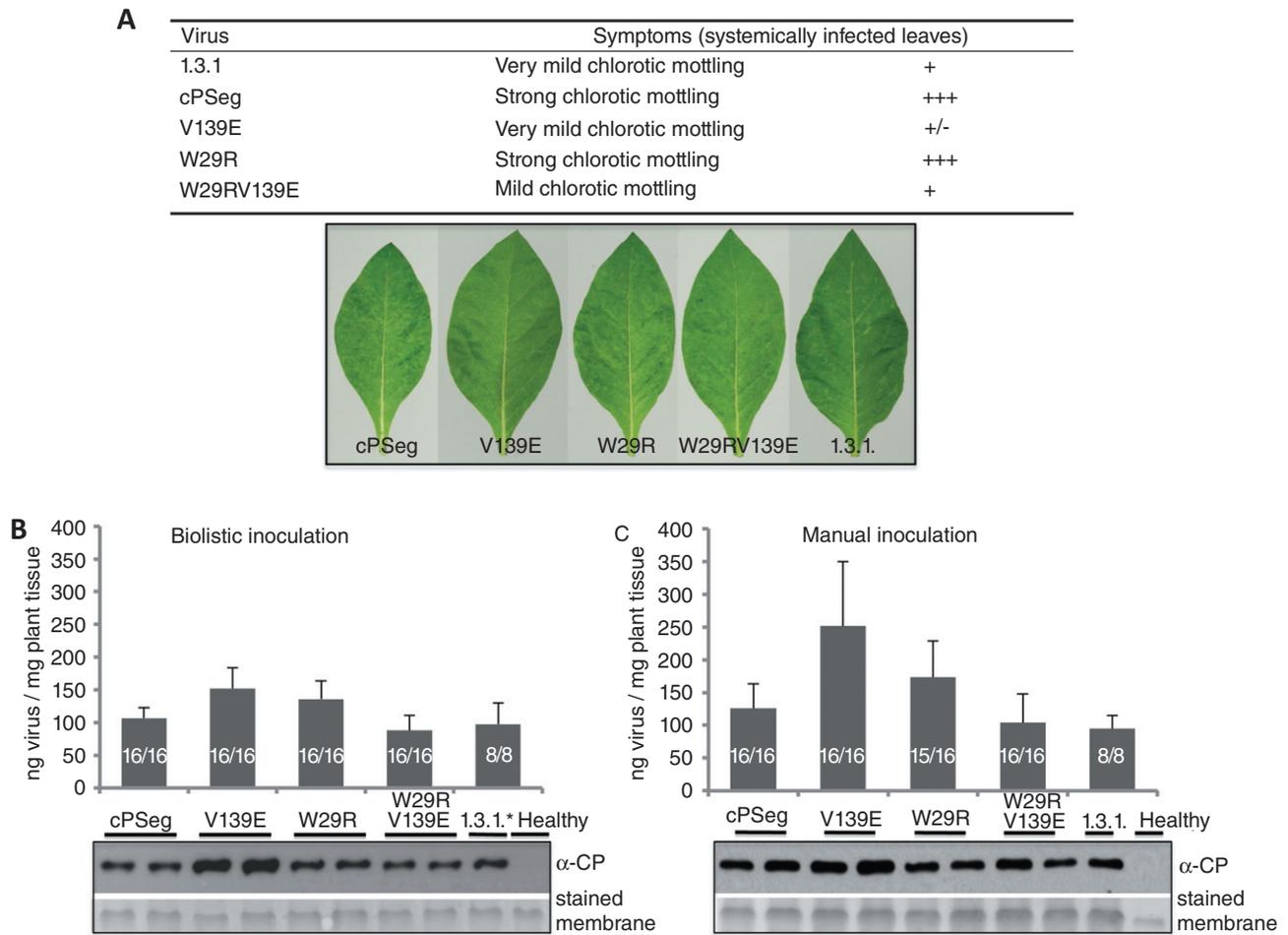


Fig. 1 Infectivity of *Plum pox virus* (PPV) isolate, PPV-PS, mutants in *Nicotiana clelandii*. (A) Photographs of systemically infected leaves of plants inoculated with the indicated mutants or the subisolate PPV-PS 1.3.1 taken at 21 days post-inoculation (dpi). The severity of the symptoms is ranked from barely detectable (+/-) to the most intense chlorotic mottling (+++). (B, C) Virus accumulation in young systemically infected leaves of *N. clelandii* biolistically (B) or manually (C) inoculated was determined by double-antibody-sandwich indirect enzyme-linked immunosorbent assay (DASI-ELISA) and Western blot at 21 dpi. In (B), the 1.3.1 subisolate was manually inoculated as no infectious cDNA was available. Bars in (B) and (C) represent average values and standard deviations from eight different plants of a representative experiment. The ratio of infected plants to the number of inoculated plants from two independent experiments is indicated inside the bars. α -CP, α -coat protein.

severe symptoms similar to those induced by PSeg, inoculation with PSegV139E caused a mild infection similar to or even milder than that caused by the natural subisolate PPV-PS 1.3.1 (Fig. 1A). Similar mild symptoms were obtained when plants were manually inoculated with leaf extracts from the infected plants initially inoculated by particle bombardment with PSegV139E. However, the levels of virus accumulation in the plants infected with PSegV139E were similar to or even higher than the virus levels of plants infected with PSeg or PSegW29R with severe symptoms (Fig. 1B,C). To test the effects of both PPV-PS 1.3.1-specific P1 changes in a PSeg background, the clone PSegW29RV139E was constructed. Surprisingly, the symptoms observed in PSegW29RV139E-infected *N. clelandii* plants were slightly, but reproducibly, more prominent than those caused by PSegV139E,

but were milder than those induced by PSeg and PSegW29R (Fig. 1A). Viral accumulation of PSegW29RV139E was similar to that of the parental PSeg (Fig. 1B,C), further confirming that the W29R and V139E P1 mutations do not have a relevant effect on virus amplification.

Virus variants segregated from the original PPV-PS isolate also differed strongly in the symptoms induced in another herbaceous host, *N. occidentalis*. In this host, severe subisolates, such as PPV-PS 4.1.4 and 10₇, or viral progeny of the clone PSeg, caused very conspicuous local necrotic lesions, which were not able to prevent the systemic spread of the virus (Sáenz *et al.*, 2001; Fig. 2A). In contrast, mild subisolates, such as PPV-PS 1.3.1, caused systemic infections without apparent symptoms in the inoculated leaves (Sáenz *et al.*, 2001; Fig. 2). Necrotic lesions similar to those

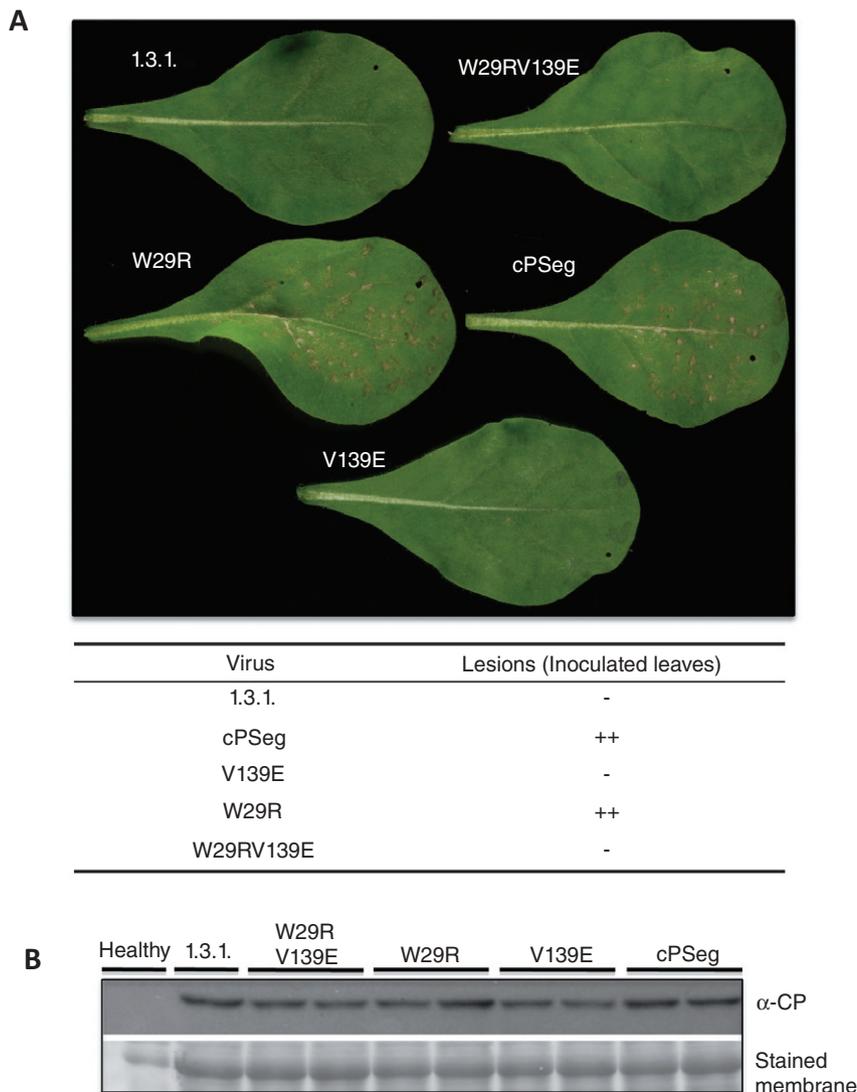


Fig. 2 Infectivity of *Plum pox virus* (PPV) isolate, PPV-PS, mutants in *Nicotiana occidentalis*. (A) Photographs of inoculated leaves taken at 10 days post-inoculation (dpi) are shown. Symptoms induced by the different viruses are described below the photographs. The intensity of the local lesions is ranked from not detectable (–) to very prominent (++) (B) Virus accumulation in manually inoculated leaves of *N. occidentalis* plants (10 dpi) was determined by Western blot. α -CP, α -coat protein.

induced by PSeg were also observed in *N. occidentalis* leaves inoculated with the PSegW29R mutant, but not in those inoculated with either the PSegV139E or PSegW29RV139E mutant (Fig. 2A). The absence of lesions in the inoculated leaves was not correlated with lower virus accumulation (Fig. 2B). Taken together, these results indicate that a glutamic acid at position 139 of the P1 protein is an attenuating factor for PPV-PS infection in both *N. clevelandii* and *N. occidentalis*. In addition, the amino acid at position 29 of the P1 protein may modulate the effect of amino acid 139 on viral pathogenesis in *N. clevelandii*.

The amino acid changes W29R and V139E affect viral infectivity and accumulation in GF305 peach seedlings

PPV-PS was originally isolated from a naturally infected peach tree. However, some of the virus variants that were isolated after propagation in *N. clevelandii* and local lesion cloning in

C. foetidum, such as PPV-PS 1.3.1, were unable to infect peach seedlings (Sáenz *et al.*, 2001). In order to assess the relevance of the amino acids present at positions 29 and 139 of the P1 protein in PPV pathogenicity in its natural host, peach GF305 seedlings were biolistically inoculated with PSeg and the mutants derived from it. PSeg and the PSegV139E mutant showed high infectivity in peach GF305 seedlings, 100% and 93.7%, respectively (Fig. 3B). Nevertheless, PSegV139E caused slightly milder symptoms and accumulated in smaller amounts than PSeg (Fig. 3). The infectivity of PSegW29R in peach seedlings was also high (85.7%), but the level of virus accumulation and symptom severity were notably lower in plants infected with this mutant than in PSeg- and PSegV139E-infected plants. Interestingly, V139E and W29R mutations had additive or synergistic effects, and the phenotype of the W29RV139E double mutant resembled that of the PPV-PS 1.3.1 subisolate, as it was noninfectious in peach GF305 seedlings (Fig. 3). Thus, although our results show that amino acid 29 of the

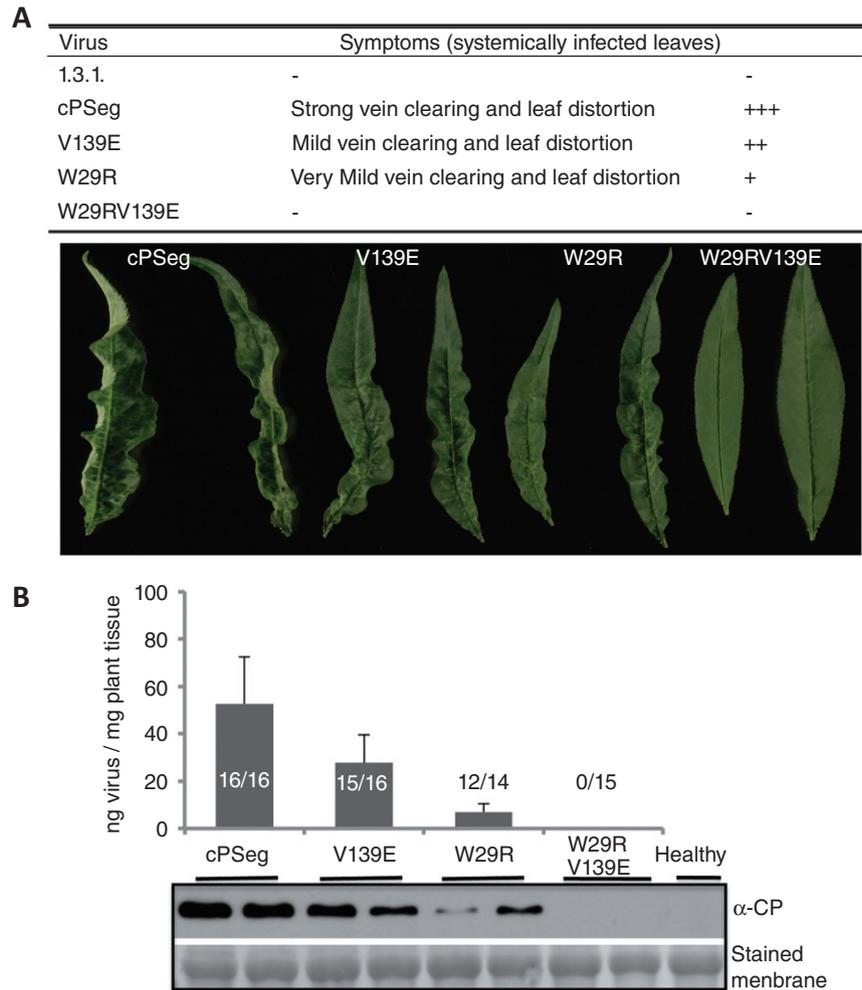


Fig. 3 Infectivity of *Plum pox virus* (PPV) isolate, PPV-PS, mutants in peach GF305 seedlings. (A) Photographs of systemically infected leaves taken at 35 days post-inoculation (dpi). The severity of the symptoms is ranked from not detectable (–) to the most intense with vein clearing and leaf distortion (+++). (B) Virus accumulation in young systemically infected leaves of biologically inoculated peach GF305 seedlings was determined by double-antibody-sandwich indirect enzyme-linked immunosorbent assay (DASI-ELISA) and Western blot at 35 dpi. Bars represent average values and standard deviations from eight different plants of a representative experiment. The ratio of infected plants to the number of inoculated plants from two independent experiments is indicated inside or above the bars. α -CP, α -coat protein.

P1 protein is particularly relevant for peach infection, similar to the important role of amino acid 139 in *Nicotiana* infection, both amino acids appear to work in concert in these two hosts.

Origin of the V139E and W29R mutations

Interestingly, *in silico* analyses of the abundant PPV sequences deposited in public databases showed that W29 and V139 were conserved in all PPV isolates, belonging to seven different strains, whose P1 sequence has been determined (Fig. S1, see Supporting Information). This suggests that there is a strong selection pressure to maintain W29 and V139, and raises the question of where virus variants with R29 and E139 emerged.

Sequence comparison of the five PPV-PS subsolates under study showed a clear clustering of severe and mild subsolates (Table 1), thus arguing against the existence of a single quasispecies cloud in the original PPV-PS isolate from which distinct subsolates with specific mutations were selected during the local lesion passage. Thus, the most probable explanation suggests the coexistence of discrete PPV variants in the original PPV-PS isolate.

As the natural peach PPV-PS isolate is no longer available, the possibility that virus variants with W29R and V139E mutations already existed in the infected tree cannot be tested. Although the infected *N. clevelandii* tissue used for the local lesion passage has not been conserved, the original PPV-PS isolate has been further propagated in *N. clevelandii*, and infected tissue of a later passage is available. Thus, deep sequencing analysis of small RNAs of this tissue was conducted using the Solexa Illumina system to identify minor sequence variants in the virus population. The Solexa Illumina sequencing technology was used because it provides a greater number of sequencing reads than other technologies available to us, and the depth of coverage allowed us to assemble complete genomes and to assess the sequence diversity present in a tissue sample. About 4.5×10^6 reads were obtained, and 62.67% of them were identified as being derived from PPV. They cover the complete viral genome (Fig. S2, see Supporting Information), allowing us to reassemble the complete PPV-PS sequence. Because each nucleotide position was covered by a large number of small interfering RNA (siRNA) reads, the siRNA library also allowed us to assess the sequence heterogeneity of the

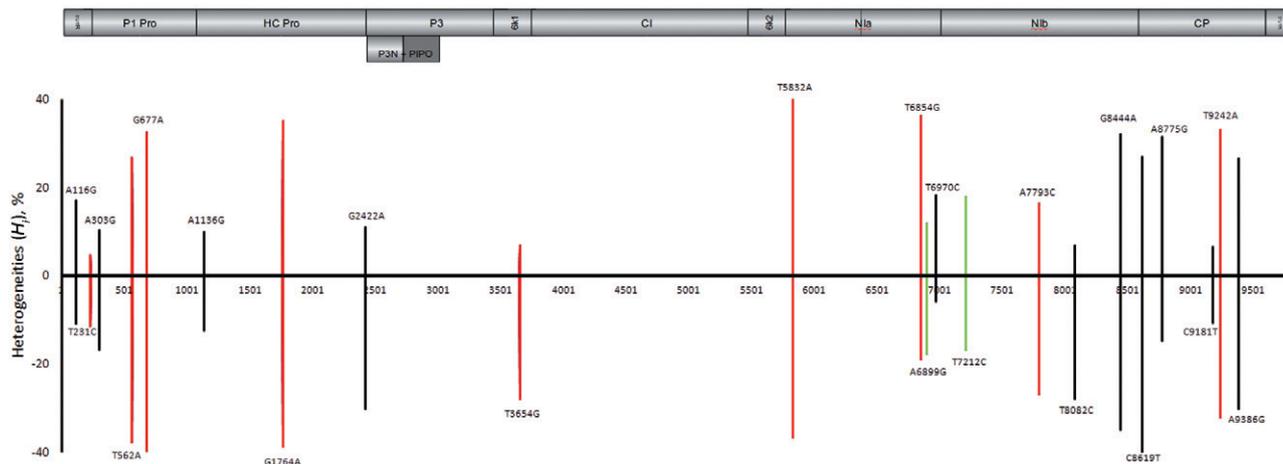


Fig. 4 Heterogeneity map of the *Plum pox virus* (PPV) isolate, PPV-PS, population. The height of the lines represents the heterogeneity values (H_i) of the indicated position for small interfering RNAs (siRNAs) that match the viral plus (up lines) or minus (down lines) strand. Only values of positions covered by more than 40 siRNA reads and more than 3% of heterogeneity for each strand are shown. Red and green lines correspond to heterogeneities segregated in PS subisolates that correlate with symptom severity (red lines) or differentiate PS10; from the rest of the subisolates (green lines). Black lines correspond to heterogeneities that do not segregate in PS subisolates. Labels show the two most abundant nucleotides at the indicated position. A schematic representation of the PPV genome is shown above.

population. In order to avoid false heterogeneities derived from cloning and sequencing artefacts, we only considered nucleotide changes that were detected in both viral reads of plus and minus polarity.

Interestingly, heterogeneities did not form the uniform cloud of mutations that would be expected from a single virus quasispecies, but discrete heterogeneity hot spots were detected (Fig. 4), suggesting that delimited virus variants coexisted in the original PPV-PS isolate.

Most of the heterogeneities associated with the clustering of mild and severe isolates, including T231C and T562A yielding W29R and V139E changes in P1, were detected in the siRNA library (red lines in Fig. 4), which is in agreement with the assumption that virus variants quite similar to those segregated by local lesion passages already coexisted in the original PPV-PS population. G6200A was the only heterogeneity that cosegregated with mild and severe isolates which was not observed in the uncloned PPV-PS isolate. The most probable explanation for this exception is that mild-type variants of the original PPV-PS isolate had A at position 6200 at the time of the local lesion cloning, but a shift to G has taken place during propagation in *N. clelandii* sometime since then and the construction of the siRNA library.

Some heterogeneities detected in the siRNA library of uncloned PPV-PS were not represented in any of the PPV-PS subisolates (black lines in Fig. 4). This could be a result of the existence in the uncloned PPV-PS population of additional virus variants that were not selected in the clonal passages, either by chance or because they were not able to live alone in *C. foetidum*, but other explanations are also possible.

Most point mutations that appear in single subisolates, and are thus not associated with symptom severity, were not detected in the deep sequencing analysis of the original virus population (compare Table 1 and Fig. 4), probably because they were imprinted in specific lineages during the local lesion passages or during the subsequent propagation in *N. clelandii*. A couple of interesting exceptions are G6899 and C7212, which are specific to PPV-PS10, but are clearly detected in the siRNA library of uncloned PPV-PS (green lines in Fig. 4). The fact that PPV-PS10, shows a hypervirulent phenotype clearly distinguishable from that of the other severe subisolates, PPV-PS 4.1.4 and PPV-PS 5.1.3, which have the habitual A6899 and T7212 nucleotides, suggests that PPV-PS10, existed in the complex PPV-PS isolate as a third independent variant, together with the variants represented by PPV-PS 4.1.4 and PPV-PS 5.1.3 on the one hand, and PPV-PS 1.3.1 and PPV-PS 2.1.1 on the other.

DISCUSSION

Single amino acid changes in the HCPro protein between different PPV-PS subisolates have a drastic effect on viral symptoms in herbaceous hosts and on virus infectivity in peach seedlings (Sáenz *et al.*, 2001). In this work, we identify two amino acid changes in the P1 protein that affect the pathogenicity of these subisolates in a host-specific manner. The function of the potyviral P1 is almost unknown (Rohozkova and Navratil, 2011). Previous characterization of the P1 protein includes analyses of its proteinase activity (Verchot and Carrington, 1995a; Verchot *et al.*, 1991), ability to bind single-stranded RNA (Brantley and Hunt, 1993; Soumounou and Laliberté, 1994), role in genome

amplification (Verchot and Carrington, 1995b), capacity to enhance the silencing suppression activity of the HCPro protein (Pruss *et al.*, 1997; Rajamäki *et al.*, 2005; Valli *et al.*, 2006) and involvement in overcoming eukaryotic translation initiation factor 4E (eIF4E)-mediated resistance (Nakahara *et al.*, 2010). Based on sequence analysis, it has been hypothesized that the P1 protein has a role in adaptation to the host of potyviruses (Adams *et al.*, 2005; Valli *et al.*, 2007). In support of this hypothesis, nucleotide changes in the P1 sequence were associated with PPV adaptation and symptom modulation in *N. clevelandii* (Salvador *et al.*, 2008a), and with attenuation of *Papaya ringspot virus* in papaya (Chiang *et al.*, 2007). Furthermore, substitution of the PPV P1 cistron by the corresponding region of the *Tobacco vein mottling virus* (TVMV) genome did not affect PPV infectivity in common PPV/TVMV hosts, but did prevent infection of peach, which is not a host for TVMV (Salvador *et al.*, 2008b). The difference in PPV-PS pathogenicity in *Nicotiana* and *Prunus*, based on the presence of V139E and W29R substitutions, further supports the existence of species-specific interactions between P1 and host factors that influence host-specific virus infectivity. Although our data do not address the details of these interactions, the fact that changes in both P1 and HCPro affect the pathogenicity of PPV-PS subisolates suggests that the ability of P1 to enhance the silencing suppression activity of HCPro may contribute, directly or indirectly, to the role of P1 in host adaptation and species-specific pathogenicity. However, the involvement of P1 and HCPro in two independent mechanisms affecting virus pathogenesis cannot be ruled out.

The isolation of variants from the original PPV-PS isolate propagated in *N. clevelandii* indicates that the PS isolate consists of a complex virus population (Sáenz *et al.*, 2001). The low degree of divergence between different PPV-PS subisolates could agree with a quasispecies distribution of the PPV-PS population. However, as discussed previously, the amino acid changes observed in the PPV-PS subisolates do not appear to follow a random distribution profile expected for a single quasispecies (Sáenz *et al.*, 2001). The deep sequencing analysis of the uncloned PPV-PS isolate further supports this conclusion. This analysis does not allow us to discriminate between multiple variants with few heterogeneities or a reduced number of variants with more nucleotide changes. However, although PPV-PS 1.3.1, 2.1.1, 4.1.4 and 5.1.3 were derived from four independent local lesions of *C. foetidum*, and were cloned through two additional local lesion passages in this host and then propagated in *N. clevelandii*, PPV-PS 1.3.1 and 2.1.1 differed in only two nucleotides and PPV-PS 4.1.4 and 5.1.3 differed in three nucleotides, whereas there were nine nucleotide changes shared by these pairs of isolates, all easily detected in the deep sequencing analysis of the uncloned PPV-PS isolate. This suggests that the original PPV-PS population consisted of a limited number of virus variants, and these four subisolates derived from two of them, each with particular host-dependent pathogenicities.

The severe subisolate PPV-PS 10₇ has some pathogenicity peculiarities. It shows a highly severe phenotype in *N. clevelandii*, but appears to infect peach GF305 seedlings less efficiently than PPV-PS 4.1.4 and 5.1.3 (Sáenz *et al.*, 2001). Although it shares with the other two severe isolates 10 nucleotides that differentiate them from the mild isolates, there are six nucleotides specific to PPV-PS 10₇, and two can be recognized in the deep sequencing library of uncloned PPV-PS. Thus, PPV-PS 10₇ could be representative of a third virus variant coexisting in the original PPV-PS. Interestingly, the previously reported cloned PS-MCI appears to be a chimera derived from different virus variants, which is in agreement with the fact that it was constructed by assembly of partial cDNA fragments cloned directly from the original PPV-PS isolate (Sáenz *et al.*, 2000).

Evidence for the existence of complex viral populations evolving independently in a woody plant has been reported for PPV (Jridi *et al.*, 2006). However, the possibility that the virus variant represented by the subisolates PPV-PS 1.3.1 and 2.1.1, with P1 R29E139, was already present in the original naturally infected tree appears to be quite unlikely. A virus variant expressing a P1 protein containing R29 and E139 would not be able to systemically infect peach, and W29 and V139 are conserved in all PPV P1 sequences available in public databases. As the P1 protein has been shown to be active *in trans* (Verchot and Carrington, 1995b), it is possible that a virus providing helper P1 activity could allow an R29E139 variant to survive in a mixed PPV population. However, the fact that only PPV-PSeg could be rescued after experimental co-inoculation of PPV-PSeg and PPV-PSegW29RV139E in peach GF305 seedlings, whereas both viruses coexisted in co-inoculated *N. clevelandii* plants (data not shown), argue against this possibility, and suggest that the R29E139 variant was most probably generated during the propagation of PPV-PS in *N. clevelandii*.

The observation that heterogeneities recognized in the uncloned PPV-PS isolate by deep sequencing are more abundant than those detected in the cloned isolates suggests that either the complexity of the PPV population has increased from the time at which the local lesion passages were carried out and the timing of deep sequencing, or that the PPV-PS isolate contained virus variants that were not fit to multiply in *C. foetidum*, or were unable to survive by themselves without the collaboration of helper virus variants. In both cases, the varied repertoire of the PPV population should facilitate the virus to adapt to new environments or hosts, and variability in the P1 protein plays an important role in this process.

EXPERIMENTAL PROCEDURES

Sequencing analysis

Leaf extracts from infected *N. clevelandii* plants were homogenized in 5 mM sodium phosphate buffer (SPB), pH 7.5 (2 mL/g tissue), and

incubated in tubes previously coated with anti-PPV immunoglobulin Gs (IgGs) overnight at 4 °C. An additional incubation at 37 °C for 2 h was followed by two washing steps with phosphate-buffered saline (PBS)–Tween buffer (16 mM SPB, 0.1 M NaCl, 0.5 g/L Tween 20, pH 7.2). RT-PCR was performed using a Titan kit (Roche Molecular Biochemical, Mannheim, Germany) with primers designed from available PPV sequence data. cDNA fragments comprising the complete PPV genome were amplified. For sequence analysis, PCR fragments were purified using a Minielute PCR purification kit (Qiagen, Hilden, Germany) and sequenced on an Abi Prism 3700 machine (Applied Biosystems, Foster City, CA, USA). GeneScan software (Applied Biosystems) was used to analyse the fragments obtained.

Construction of full-length cDNA clones

To obtain the plasmid named pICPPV-PsEs, two PPV cDNA fragments were amplified from pGPPV-PsEs (Sáenz *et al.*, 2001) by PCR using primers 5'-CAGAACTCGGAATGC-3' (primer 270; nucleotides 2260–2275) and 5'-TCCTGCAGATAACTTTTTCAACCAG-3' (nucleotides 2926–2901, with a G to C substitution indicated in bold italic, creating a *Pst*I site in italic) for one of the fragments, and 5'-ATCTGCAGGAATTGGAGCAAGC-3' (nucleotides 2917–2938, with a C to G substitution indicated in bold italic, and the introduced *Pst*I site in italic) and 5'-CGAACCAACGCCACTG-3' (primer 237; nucleotides 4945–4930) for the second. These two fragments were used as templates for a new PCR amplification with primers 270 and 237, and the product was cloned into pGEM-T to generate pGEM-PsEs. A DNA fragment containing intron I from the *ST-LS-1* gene of potato (Vancanneyt *et al.*, 1990) was PCR amplified from the pGUS-intron plasmid as described previously (López-Moya and García, 2000). This PCR-amplified fragment was digested with *Pst*I and *Nsi*I, and cloned into the engineered *Pst*I site of the PPV sequence of pGEM-PsEs, giving rise to pGEM-PsEs-STLS1. In the next step, a cDNA fragment of the 5' terminal region of the PPV-PS genome was amplified by PCR from pGPPV-PsEs using primers 5'-AAAATATAAAACTCAACAC-3' (primer 29, nucleotides 1–24) and 5'-TGAACCACTATTGAACAG-3' (primer 317, nucleotides 2609–2592), and cloned into the *Stul* site of p35S_{Se}NOSB (López-Moya and García, 2000), between the CaMV 35S promoter and the NOS terminator sequences, rendering p35S5'PSNOSB. p35SNBSNOSB was obtained by inserting *Ndel*-*Bgl*II (PPV nucleotides 309–2312) and *Bgl*II-*Sal*I (PPV nucleotides 2312–7633) fragments from pGPPV-PsEs into *Sal*I/*Ndel*-digested p35S5'PSNOSB. Finally, a *Bgl*II-*Asp*I fragment from pGEM-PsEs-STLS1 (PPV nucleotides 2312–4709) was ligated into *Bgl*I-*Bgl*II (inside the vector to PPV nucleotide 2312) and *Asp*I-*Bgl*I (PPV nucleotide 4709 to inside the vector) fragments from p35SNBSNOSB to obtain the complete pICPPV-PsEs clone.

A T231C mutation to generate the W29R substitution was introduced into pICPPV-PsEs by mutagenic PCR. PPV DNA fragments 1–238 and 238–2608 were amplified by PCR from pICPPV-PsEs with primers 29 and 5'-GCAAAGCCGGGACCCG-3' (nucleotides 222–238), and 5'-GGGTCCCGCCTTTCGCG-3' (nucleotides 221–237) and 317 (the mutated positions are indicated in bold italic), respectively. The PCR products were used as templates for PCR amplification with primers 29 and 317 to obtain the PPV-PsEs fragment 1–2608 containing the T231C mutation, which was cloned into *Stul*-digested p35SNOSB plasmid giving rise to p35S5'W29RNOSB. pICPPV-PsEsW29R was finally obtained by ligating an *Nae*I-*Nde*I fragment (inside the vector to PPV nucleotide 309) of

p35S5'W29RNOSB with *Nde*I-*Bgl*II (PPV nucleotides 309–2312) and *Bgl*II-*Nae*I (PPV nucleotide 2312 to inside the vector) fragments from pICPPV-PsEs.

As a first step to introduce the T562A mutation, which causes the amino acid change V139E, in a pICPPV clone, a cDNA fragment spanning nucleotides 1–2608 was amplified by IC-RT-PCR from extracts of *N. cleavelandii* plants infected with the subsolate PPV-PS 1.3.1. The PCR product was cloned into *Stul*-digested p35SNOSB to obtain p35S5'1.3.1NOSB carrying nucleotide changes T231C, T562A and A1764G. Finally, ligation of the *Nde*I-*Bgl*II fragment (PPV nucleotides 309–2312) from p35S5'1.3.1NOSB with *Nae*I-*Nde*I (from inside the vector to PPV nucleotide 309) and *Bgl*II-*Nae*I (PPV nucleotide 2312 to inside the vector) fragments from pICPPV-PsEs yielded pICPPV-PSegV139E. This plasmid differs from pICPPV-PsEs by the nucleotide changes T562A and A764G, which give rise to the V139E and S232G mutations in the P1 and HCPro proteins, respectively.

pICPPV-PSegW29R was constructed by ligating the *Nde*I-*Bgl*II fragment (PPV nucleotides 309–2312) from pGPPV-PSeg (Sáenz *et al.*, 2001) and the *Bgl*II-*Sal*I fragment (PPV nucleotides 2312–7633) from pICPPV-PS into *Sal*I/*Nde*I-digested pICPPV-PsEsW29R.

To generate pICPPV-PSeg without P1 mutations, the *Bgl*II-*Sal*I (PPV nucleotides 2312–7633) and *Sal*I-*Nde*I (PPV nucleotides 7633–309) fragments from pICPPV-PsEs were ligated with the *Nde*I-*Bgl*II (PPV nucleotides 309–2312) fragment from pICPPV-PSegW29R.

The double P1 mutant pICPPV-PSegW29RV139E was obtained by ligating the *Nde*I-*Bgl*II (PPV nucleotides 309–2312) and *Bgl*II-*Sal*I (PPV nucleotides 2312–7633) fragments from pICPPV-PSegV139E into *Sal*I/*Nde*I-digested pICPPV-PsEsW29R.

The accuracy of all the constructs was verified by restriction digestion analysis and DNA sequencing of all regions derived from PCR amplification.

Inoculation and protein analysis

Nicotiana cleavelandii plants and peach (*P. persica*) cv. GF305 seedlings were biolistically inoculated using the Helios Gene Gun System (Bio-Rad, Hercules, CA, USA). Microcarrier cartridges were prepared from two different clones per construct, with 1.0-µm gold particles coated with pICPPV-derived plasmids at a DNA loading ratio of 2 µg/mg gold and a microcarrier loading quantity of 0.5 mg/shooting, according to the manufacturer's instructions. Helium pressures of 7.5 bar and 10 bar were used for inoculations of *N. cleavelandii* and peach GF305, respectively. Each cartridge was shot twice onto two leaves of each plant and, for each plant, one or two cartridges were administered for *N. cleavelandii* and GF305, respectively.

For manual inoculation, infected plant leaves, coming from two biolistically inoculated *N. cleavelandii* plants, or from *N. cleavelandii* leaves infected with PPV-PS 1.3.1 stored at –20 °C, were ground in 5 mM sodium phosphate (pH 7.5) with an ice-cold pestle (2 mL/g tissue). Extracts were centrifuged to eliminate tissue fragments. For each extract, three plant leaves were dusted with carborundum and inoculated with a total of 15 µL of extract.

Plants were maintained in a glasshouse with 16 h of light by supplementary illumination and kept between 19 and 23 °C. Alternatively, plants were kept in a climate-controlled chamber (22 or 16 °C) with 14 h of light. Virus infection was monitored by observation of symptoms, and images

were recorded with a Nikon D1X digital camera (Nikon Instruments Europe B.V.). Virus accumulation was assessed at various time points post-inoculation using double-antibody-sandwich indirect enzyme-linked immunosorbent assays (DASI-ELISA) with a REALISA kit (Durviz, Paterna, Valencia, Spain) and Western blot analysis.

Deep sequencing analysis

Total RNA was extracted from 0.5 g of PS-infected *N. clevelandii* leaves using 1 mL of Trizol reagent (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. Deep sequencing was performed by Progenika Biopharma, S.A. (Derio, Spain) using the Illumina Solexa platform.

Raw Illumina reads were parsed with UEA sRNA Toolkit (<http://srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi>) (Moxon *et al.*, 2008). Parsing consisted of the removal of the 3' adaptor sequence (reads with no adaptor were discarded), and the filtering of sRNA and tRNA sequences. Sequences with less than 18 nucleotides and more than 25 nucleotides were not considered in the following analyses. Processed reads were aligned against the PS-MCI genome with BWA (<http://bio-bwa.sourceforge.net/>) (Li and Durbin, 2010) allowing one mismatch without gaps.

For the heterogeneity analysis, heterogeneity was defined as $H_i = (n_i / \text{coverage}_i) \times 100$, where H_i is the percentage of heterogeneity in position i , n_i is the number of aligned siRNAs which differ from the reference in position i and coverage_i is the total number of siRNAs aligned in position i .

In-house-developed R and PHP scripts were used for all statistical analyses and for plotting the alignments, respectively.

ACKNOWLEDGEMENTS

We wish to thank Elvira Dominguez for technical assistance. This work was supported by grants BIO2010-18541 from the Spanish Ministerio de Educación y Ciencia (MEC), SAL/0185/2006 from Comunidad de Madrid and KBBE-204429 from the European Union. B.S. was a recipient of a Formación de Personal Investigador fellowship from MEC.

REFERENCES

Adams, M.J., Antoniw, J.F. and Fauquet, C.M. (2005) Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Arch. Virol.* **150**, 459–479.

Ayme, V., Petit-Pierre, J., Souche, S., Palloix, A. and Moury, B. (2007) Molecular dissection of the potato virus Y VPg virulence factor reveals complex adaptations to the pvr2 resistance allele series in pepper. *J. Gen. Virol.* **88**, 1594–1601.

Biebricher, C.K. and Eigen, M. (2006) What is a quasispecies? *Curr. Top. Microbiol. Immunol.* **299**, 1–31.

Brantley, J.D. and Hunt, A.G. (1993) The N-terminal protein of the polyprotein encoded by the potyvirus tobacco vein mottling virus is an RNA-binding protein. *J. Gen. Virol.* **74**, 1157–1162.

Charron, C., Nicolai, M., Gallois, J.L., Robaglia, C., Moury, B., Palloix, A. and Caranta, C. (2008) Natural variation and functional analyses provide evidence for co-evolution between plant eIF4E and potyviral VPg. *Plant J.* **54**, 56–68.

Chiang, C.H., Lee, C.Y., Wang, C.H., Jan, F.J., Lin, S.S., Chen, T.C., Raja, J.A.J. and Yeh, S.-D. (2007) Genetic analysis of an attenuated *Papaya ringspot virus* strain applied for cross-protection. *Eur. J. Plant Pathol.* **118**, 333–348.

Chung, B.Y.W., Miller, W.A., Atkins, J.F. and Firth, A.E. (2008) An overlapping essential gene in the *Potyviridae*. *Proc. Natl. Acad. Sci. USA*, **105**, 5897–5902.

Domingo, E. and Holland, J.J. (1997) RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* **51**, 151–178.

Eigen, M. (1996) On the nature of virus quasispecies. *Trends Microbiol.* **4**, 216–218.

García, J.A. and Cambra, M. (2007) Plum pox virus and sharka disease. *Plant Viruses*, **1**, 69–79.

Hajimorad, M.R., Wen, R.H., Eggenberger, A.L., Hill, J.H. and Maroof, M.A. (2011) Experimental adaptation of an RNA virus mimics natural evolution. *J. Virol.* **85**, 2557–2564.

Holmes, E.C. and Moya, A. (2002) Is the quasispecies concept relevant to RNA viruses? *J. Virol.* **76**, 460–465.

Jenkins, G.M., Worobey, M., Woelk, C.H. and Holmes, E.C. (2001) Evidence for the non-quasispecies evolution of RNA viruses. *Mol. Biol. Evol.* **18**, 987–994.

Jridi, C., Martin, J.F., Marie-Jeanne, V., Labonne, G. and Blanc, S. (2006) Distinct viral populations differentiate and evolve independently in a single perennial host plant. *J. Virol.* **80**, 2349–2357.

Li, H. and Durbin, R. (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, **26**, 589–595.

López-Moya, J.J. and García, J.A. (2000) Construction of a stable and highly infectious intron-containing cDNA clone of plum pox potyvirus and its use to infect plants by particle bombardment. *Virus Res.* **68**, 99–107.

Malpica, J.M., Fraile, A., Moreno, I., Obies, C.I., Drake, J.W. and Garcia-Arenal, F. (2002) The rate and character of spontaneous mutation in an RNA virus. *Genetics*, **162**, 1505–1511.

Moxon, S., Schwach, F., MacLean, D., Dalmay, T., Studholme, D.J. and Moulton, V. (2008) A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatics*, **24**, 2252–2253.

Nakahara, K.S., Shimada, R., Choi, S.-H., Yamamoto, H., Shao, J. and Uyeda, I. (2010) Involvement of the P1 cistron in overcoming eIF4E-mediated recessive resistance against *Clover yellow vein virus* in pea. *Mol. Plant–Microbe Interact.* **23**, 1460–1469.

Ohshima, K., Akaiishi, S., Kajiyama, H., Koga, R. and Gibbs, A.J. (2010) Evolutionary trajectory of turnip mosaic virus populations adapting to a new host. *J. Gen. Virol.* **91**, 788–801.

Pruss, G., Ge, X., Shi, X.M., Carrington, J.C. and Vance, V.B. (1997) Plant viral synergism: the potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell*, **9**, 859–868.

Rajamäki, M.L., Kelloniemi, J., Alminaita, A., Kekarainen, T., Rabenstein, F. and Valkonen, J.P. (2005) A novel insertion site inside the potyvirus P1 cistron allows expression of heterologous proteins and suggests some P1 functions. *Virology*, **342**, 88–101.

Rohozkova, J. and Navratil, M. (2011) P1 peptidase—a mysterious protein of family *Potyviridae*. *J. Biosci.* **36**, 189–200.

Sáenz, P., Cervera, M.T., Dallot, S., Quiot, L., Quiot, J.B., Riechmann, J.L. and García, J.A. (2000) Identification of a pathogenicity determinant of *Plum pox virus* in the sequence encoding the C-terminal region of protein P3 + 6K1. *J. Gen. Virol.* **81**, 557–566.

Sáenz, P., Quiot, L., Quiot, J.-B., Candresse, T. and García, J.A. (2001) Pathogenicity determinants in the complex virus population of a *Plum pox virus* isolate. *Mol. Plant–Microbe Interact.* **14**, 278–287.

Salvador, B., García, J.A. and Simón-Mateo, C. (2006) Causal agent of sharka disease: *plum pox virus* genome and function of gene products. *EPPO Bull.* **36**, 229–238.

Salvador, B., Delgadillo, M.O., Sáenz, P., García, J.A. and Simón-Mateo, C. (2008a) Identification of *Plum pox virus* pathogenicity determinants in herbaceous and woody hosts. *Mol. Plant–Microbe Interact.* **21**, 20–29.

Salvador, B., Sáenz, P., Yanguel, E., Quiot, J.B., Quiot, L., Delgadillo, M.O., García, J.A. and Simón-Mateo, C. (2008b) Host-specific effect of P1 exchange between two potyviruses. *Mol. Plant Pathol.* **9**, 147–155.

Soumounou, Y. and Laliberté, J.-F. (1994) Nucleic acid-binding properties of the P1 protein of turnip mosaic potyvirus produced in *Escherichia coli*. *J. Gen. Virol.* **75**, 2567–2573.

Suehiro, N., Natsuaki, T., Watanabe, T. and Okuda, S. (2004) An important determinant of the ability of *Turnip mosaic virus* to infect *Brassica* spp. and/or *Raphanus sativus* is in its P3 protein. *J. Gen. Virol.* **85**, 2087–2098.

Valli, A., Martín-Hernández, A.M., López-Moya, J.J. and García, J.A. (2006) RNA silencing suppression by a second copy of the P1 serine protease of *Cucumber vein yellowing ipomovirus* (CVYV), a member of the family *Potyviridae* that lacks the cysteine protease HCPro. *J. Virol.* **80**, 10 055–10 063.

Valli, A., López-Moya, J.J. and García, J.A. (2007) Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family *Potyviridae*. *J. Gen. Virol.* **88**, 1016–1028.

Vancanneyt, G., Schmidt, R., O'Connor-Sanchez, A., Willmitzer, L. and Rocha-Sosa, M. (1990) Construction of an intron-containing marker gene: splicing of the intron in transgenic plants and its use in monitoring early events in Agrobacterium-mediated plant transformation. *Mol. Gen. Genet.* **220**, 245–250.

- Verchot, J. and Carrington, J.C. (1995a) Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications. *J. Virol.* **69**, 1582–1590.
- Verchot, J. and Carrington, J.C. (1995b) Evidence that the potyvirus P1 proteinase functions in trans as an accessory factor for genome amplification. *J. Virol.* **69**, 3668–3674.
- Verchot, J., Koonin, E.V. and Carrington, J.C. (1991) The 35-kDa protein from the N-terminus of a potyviral polyprotein functions as a third virus-encoded proteinase. *Virology*, **185**, 527–535.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Fig. S1 Multiple amino acid alignment of the P1 proteins of 97 isolates belonging to seven *Plum pox virus* strains recovered from public databases. The amino acid conservation at positions W29 and V139E is highlighted.

Fig. S2 Alignment of small RNAs to the reference PS-MCL genome. The green bars represent the unique small RNAs (length between 20 and 25 bases) aligned to the two strands of the PS-MCL genome. Only perfect matching was allowed.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

