

Document downloaded from:

<http://hdl.handle.net/10251/161162>

This paper must be cited as:

Maliogka, VI.; Salvador, B.; Carbonell, A.; Saenz, P.; San Leon, D.; Oliveros, JC.; Delgadillo, MO.... (2012). Virus variants with differences in the P1 protein coexist in a Plum pox virus population and display particular host-dependent pathogenicity features. *Molecular Plant Pathology*. 13(8):877-886. <https://doi.org/10.1111/j.1364-3703.2012.00796.x>



The final publication is available at

<https://doi.org/10.1111/j.1364-3703.2012.00796.x>

Copyright Blackwell Publishing

Additional Information

**VIRUS VARIANTS WITH DIFFERENCES IN THE P1 PROTEIN COEXIST IN A
PLUM POX VIRUS POPULATION AND DISPLAY PARTICULAR HOST-DEPENDENT
PATHOGENICITY FEATURES**

Varvara I. Maliogka¹✉, Beatriz Salvador²✉, Alberto Carbonell³, Pilar Sáenz⁴, David San León,
Juan Carlos Oliveros, M. Otilia Delgadillo, Juan Antonio García[§], Carmen Simón-Mateo[§]

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

✉These authors contribute the same to the work

§Corresponding authors: E-mail: jagarcia@cnb.uam.es, csimon@cnb.uam.es, phone: +34
915854535, fax: +34 915854506

Other E-mail addresses:

VIM: vmaliogk@agro.auth.gr

BS: salvador@cnb.uam.es

AC: acarbonell@danforthcenter.org

PS: psaenz@progenika.com

MOD: modelgad@cnb.uam.es

¹ Present address: Plant Pathology Laboratory, School of Agriculture, Aristotle University of
Thessaloniki, 54124 Thessaloniki, Greece

² Present address:

³ Present address: Donald Danforth Plant Science Center, 975 N. Warson Rd., St Louis, MO
63132, USA.

⁴ Present address: PROGENIKA BIOPHARMA S.A., Edificio 801, Parque Tecnológico de
Zamudio, 48160 Derio, Spain

Abstract

Subisolates segregated from a 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 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 some 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.

Keywords: *Plum pox virus*, sharka, PPV, potyvirus, pathogenicity determinants, host range, quasispecies, deep sequencing

Introduction

Plum pox virus (PPV) is a potyvirus that causes sharka disease in stone fruit trees of the genus *Prunus* [1]. The PPV genome consists of a single-stranded RNA molecule, 9786 nt in length, which encodes a single polyprotein and a truncated frameshift product, which are proteolytically processed by three self-encoded proteases [2,3]. 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 [4]. This, in combination with short replication times and high RNA yields give rise to complex and dynamic mutant clouds. This population structure is usually interpreted in the terms of the quasispecies theory [5,6,7], although other conceptual models have also been proposed [8,9]. 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 one centered on a consensus sequence, may coexist in single individuals. In these cases, the pathogenicity displayed is the result of complex intra- and inter-population interactions and not necessarily a simple combination of the pathologies exhibited by the separate virus populations [5].

The PPV-PS isolate belonging to the M strain was originally obtained from a severely infected peach tree in former Yugoslavia, and propagated in *Nicotiana clevelandii*. Studies of local lesion cloning in *Chenopodium foetidum* revealed the complexity of the original virus population [10]. Several PPV-PS subisolates were segregated and found to share a high degree of sequence similarity (~99.9%) despite manifesting large differences in pathogenicity in herbaceous and woody hosts [10].

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 HCPro protein produced a drastic enhancement of symptom severity in both *N. clevelandii* and *N. 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 [10]. 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 aminoacid changes in their P1 protein

The complete genomic sequence 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 *Chenopodium foetidum* [10], was determined on partial cDNA fragments amplified by RT-PCR 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 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) [11] in 6 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 non-conservative amino acid substitutions in the P1 protein (W29R

and V139E), and one conservative amino acid substitution in the VPg protein (S40T), with respect to the PS-MCI sequence. These latter 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 sequence has been determined in this work (Table 1).

The amino acid change V139E affects symptoms development in N. clevelandii 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. Since 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 [10], 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 CaMV 35S promoter, which allows the production *in planta* of infectious transcripts after inoculation with plasmid DNA [12]. Inoculation with PSeg caused severe symptoms in *N. clevelandii* and *N. occidentalis* that were indistinguishable from those induced by the severe PPV-PS subisolates [Figs. 1 and 2, and 10]. 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. clevelandii* plants (Fig. 1). Both clones infected 100% of biolistically inoculated plants in two independent experiments. While inoculation with

PSegW29R caused 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, those infected with PSeg or PSegW29 with severe symptoms (Figs. 1 B and C). To test the effect of both PPV-PS 1.3.1-specific P1 changes in a PSeg background, the clone PSegW29RV139E was constructed. Surprisingly, symptoms observed in PSegW29RV139E-infected *N. clevelandii* plants were slightly, but reproducibly, more prominent than those caused by PSegV139E, yet they 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 and 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 largely 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 and Fig. 2A). In contrast, mild subisolates, such as PPV-PS 1.3.1, caused systemic infections without apparent symptoms in the inoculated leaves [10 and Fig. 2]. Necrotic lesions similar to those 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 mutants (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 [10]. 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 hosts, peach GF305 seedlings were biolistically inoculated with PSeg and the mutants derived from it. PSeg and the PSegV139E showed high infectivity in peach GF305 seedlings, 100% and 93.7%, respectively (Fig. 3B). Nevertheless, PSegV139E caused slightly milder symptoms and accumulated in lower amount than PSeg (Fig. 3). Infectivity of PSegW29R in peach seedlings was also high (85.7%), but the level of virus accumulation and symptom severity, were notably lower in the plants infected with this mutants than in the 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 in that it was non-infectious in peach GF305 seedlings (Fig. 3). Thus, although our results show that amino acid 29 of the 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, informatic searches on the abundant PPV sequences deposited in public databases showed that W29 and V139 were conserved in all PPV isolates, belonging to seven different PPV strains, whose P1 sequence has been determined (Supplementary Fig. 1). 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 subisolates under study showed a clear clustering of severe and mild subisolates (Table 1), thus arguing against the existence of a single quasispecies cloud in the original PPV-PS isolate from which distinct subisolates with specific mutations were selected during the local lesion passage. Thus, the most likely explanation suggests the coexistence of discrete PPV variants in the original PPV-PS isolate.

Since the natural peach PPV-PS isolate is not 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 by the Solexa Illumina system, trying to identify minor sequence variants in the virus population. 2.8×10^6 reads were obtained, and xx% of them were identified as derived from PPV. They cover the complete viral genome (Supplementary Fig. 2), allowing us to

reassemble the complete PPV-PS sequence. Because each nucleotide position was covered by a large number of siRNA reads, the siRNA library also allowed us to assess the sequence heterogeneity of the 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 were not forming the quite 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 were coexisting in the original PPV-PS isolate.

Most of the heterogeneities associated to 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 that was not observed in the uncloned PPV-PS isolate. The most likely 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. clevelandii* 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 due to 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, thus, are 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. clevelandii*. A couple of interesting exceptions are G6899 and C7212, which are specific of PPV-PS10₇, but are clearly detected in the siRNA library of uncloned PPV-PS (green lines in Fig. 4). The fact that PPV-PS10₇ shows an 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-PS10₇ 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 one hand, and PPV-PS 1.3.1 and PPV-PS 2.1.1 on the other hand.

Discussion

Single amino acid changes in the HCPro protein between different PPV-PS subisolates had a drastic effect on viral symptoms in herbaceous hosts, and on virus infectivity in peach seedlings as well [10]. 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 fairly unknown [13]. Previous characterization of the P1 protein includes analyses of its proteinase activity [14,15], ability to bind single-stranded RNA [16,17], role in genome amplification [18], capacity to enhance the silencing suppression activity of the HCPro protein [19,20,21], and involvement in overcoming eIF4E-mediated resistance [22]. Based on sequence analysis, it has been hypothesized that the P1 protein has a role in adaptation

to the host of potyviruses [23,24]. In support of this hypothesis, nucleotide changes in the P1 sequence were associated with PPV adaptation and symptom modulation in *N. clevelandii* [25], and with attenuation of *Papaya ringspot virus* in papaya [26]. 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 [27]. 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 P1 role 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 [10]. The low degree of divergence between different PPV-PS subisolates could agree with a quasispecies distribution of the PPV-PS population. However, as it has been previously discussed, the amino acid changes observed in the PPV-PS subisolates do not appear to follow a random distribution profile expected for a single quasispecies [10]. The deep sequencing analysis of the uncloned PPV-PS isolate further supports this conclusion. This analysis does not allow 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 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, 2.1.1 differ in only 2 nt and PPV-PS 4.1.4 and 5.1.3 differ in 3 nt, whereas there are 9 nt changes shared by these couples of isolates, all of them 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 derive from two of them, each one with particular host-dependent pathogenicities. The severe subisolate PPV-PS 10₇ has some pathogenicity peculiarities. It shows 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 [10]. Although it shares with the other two severe isolates 10 nt that differentiate them from the mild isolates, there are 6 nt specific of PPV-PS 10₇, and two of them can be recognized in the deep sequencing library of uncloned PPV-PS. Thus, PPV-PS 10₇ could be representative of a third virus variants coexisting in the original PPV-PS. Interestingly, the previously reported cloned PS-MCl 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 [11].

Evidence for the existence of complex viral populations evolving independently in a woody plant has been reported for PPV [28]. 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. Since the P1 protein has been shown to be active in

trans [18], it would be possible that a virus providing helper P1 activity could allow a R29E139 variant to survive in a mixed PPV population. However, the fact that only PPV-PSeg could be rescued after experimental coinoculation of PPV-PSeg and PPV-PSegW29RV139E in peach GF305 seedlings, whereas both viruses coexisted in coinoculated *N. clevelandii* plants (data not shown), argue against this possibility, and suggest that the R29E139 variant was likely generated during propagation of PPV-PS in *N. clevelandii*.

The fact 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 the local lesion passages were carried out and the timing of the deep sequencing, or that the PPV-PS isolate contained virus variants that were not fitted to multiply in *C. foetidum*, or were unable to survive by themselves without the collaboration or 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.

Materials and methods

Sequencing analysis

Leaf extracts from infected *Nicotiana 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 IgGs overnight at 4°C. An additional incubation at 37°C for 2 h was followed by two washing steps with 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) 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) and sequenced on an Abi Prism 3700 machine. GeneScan software 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 [10] by PCR using primers 5'-CAGAAACTCGGAATGC-3' (primer 270, nt 2260-2275) and 5'-**TCCTGCAG**GATAACTTTTTTCAACCAG-3' (nt 2926-2901, with a G to C substitution indicated in bold, creating a PstI site, underlined) for one of the fragments, and 5'-AT**CTGCAG**GGAATTGGAGCAAGC-3' (nt 2917-2938, with a C to G substitution indicated in bold, and the introduced PstI site underlined) and 5'-CGAACCAACGCCACTG-3' (primer 237, nt 4945-4930) for the second one. These two fragments were used as templates for a new PCR amplification with primers 270 and 237, and the product was cloned in pGEM-T to generate pGEM-PSes. A DNA fragment containing intron I from the ST-LS-1 gene of potato [29] was PCR amplified from the pGUS-intron plasmid as previously described [12]. This PCR-amplified fragment was digested with PstI and NsiI, and cloned into the engineered PstI 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-PSEes using primers 5'-AAAATATAAAAACACTCAACAC-3' (primer 29, nt 1-24) and 5'-TGAACCACTATTGAACAG-3' (primer 317, nt 2609-2592), and cloned in the StuI site of p35SeNOSB [12], between the CaMV 35S promoter and the NOS terminator

sequences, rendering p35S5'PSNOSB. p35SNBSNOSB was obtained by inserting NdeI-BglII (PPV nt 309-2312) and BglII-SalI (PPV nt 2312-7633) fragments from pGPPV-PSes into SalI/NdeI-digested p35S5'PSNOSB. Finally, a BglII-AspI fragment from pGEM-PSes-STLS1 (PPV nt 2312-4709) was ligated to BglI-BglII (inside the vector to PPV nt 2312) and AspI-BglII (PPV nt 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 in 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'-GCAAAGGCCGGGACCCG-3' (nt 222-238), and 5'-GGGTCCCGGCCTTTGCG-3' (nt 221-237) and 317 (the mutated positions are indicated in bold), 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 StuI-digested p35SNOSB plasmid giving rise to p35S5'W29RNOSB. pICPPV-PSesW29R was finally obtained by ligating a NaeI-NdeI fragment (inside the vector to PPV nt 309) of p35S5'W29RNOSB with NdeI-BglII (PPV nt 309-2312) and BglII-NaeI (PPV nt 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 nt 1-2608 was amplified by IC-RT-PCR from extracts of *N. clelandii* plants infected with the subisolate PPV-PS 1.3.1. The PCR product was cloned into StuI-digested p35SNOSB to obtain p35S5'1.3.1NOSB carrying nucleotides changes T231C and T562A and A1764G. Finally, ligation of NdeI-BglII fragment (PPV nt 309-2312) from p35S5'1.3.1NOSB with NaeI-NdeI (from inside the vector to PPV nt 309) and BglII-NaeI (PPV nt 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 NdeI-BglII fragment (PPV nt 309-2312) from pGPPV-PSeg [10] and the BglII-SalI fragment (PPV nt 2312-7633) from pICPPV-PS into SalI/NdeI-digested pICPPV-PSesW29R.

To generate pICPPV-PSeg without P1 mutations the BglII-SalI (PPV nt 2312-7633) and SalI-NdeI (PPV nt 7633-309) fragments from pICPPV-PSes were ligated with the NdeI-BglII (PPV nt 309-2312) fragment from pICPPV-PSegW29R.

The double P1 mutant pICPPV-PSegW29RV139E was obtained by ligating the NdeI-BglII (PPV nt 309-2312) and BglII-SalI (PPV nt 2312-7633) fragments from pICPPV-PSegV139E into Sal I/NdeI-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

N. clevelandii plants and peach (*Prunus persica*) cv. GF305 seedlings were biolistically inoculated using the Helios Gene Gun System (Bio-Rad, Hercules, CA). Microcarrier cartridges were prepared from 2 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. clevelandii* 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. clevelandii* and GF305, respectively.

For manual inoculation, infected plant leaves, coming from 2 biologically inoculated *N. cleavelandii* plants, or from *N. cleavelandii* leaves infected with PPV-PS 1.3.1 stored at -20°, were ground in 5 mM sodium phosphate (pH 7.5) with an ice-cold pestle (1g : 2 ml). 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 greenhouse 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 °C 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. 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) and western blot analysis.

Deep sequencing analysis

Total RNA was extracted from 0.5 g of PS-infected *N. cleavelandii* leaves using 1 ml Trizol Reagent (Invitrogen) according to the manufacturer's instructions. Deep sequencing was performed by Progenika Biopharma, S.A. using the Illumina Solexa platform.

Raw Illumina reads were preprocessed with UEA sRNA Toolkit (<http://srna-tools.cmp.uea.ac.uk/plant/cgi-bin/srna-tools.cgi>) [30]. Preprocessing consisted in removal of 3' adaptor sequence (reads with no adaptor were discarded), and the filtering of sRNA and tRNA sequences. Sequences with less than 18 nts and more than 25 nts were not considered in the following analysis. Processed reads were

aligned against PS-MCI genome with BWA (<http://bio-bwa.sourceforge.net/>) [31] allowing one mismatch without gaps.

For the herogeneity analysis, heterogeneity was defined as:

$$H_i = \frac{n_i}{coverage_i} \times 100$$

H_i : Percentage of heterogeneity in the position i .

n_i : number of aligned siRNAs which differ from the reference in the position i .

$coverage_i$: total number of siRNAs aligned in the position i .

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

Acknowledgements

We wish to thank Elvira Dominguez for technical assistance. This work was supported by grants BIO2007-67283 and BIO2010-18541 from Spanish MEC, SAL/0185/2006 from Comunidad de Madrid, and SP22-CT-2004 from European Union. B.S. was a recipient of an FPI fellowship from MEC.

References

1. García JA, Cambra M (2007) Plum pox virus and sharka disease. *Plant Viruses* 1: 69-79.
2. Chung BYW, Miller WA, Atkins JF, Firth AE (2008) An overlapping essential gene in the Potyviridae. *Proceedings of the National Academy of Sciences USA* 105: 5897-5902.
3. Salvador B, García JA, Simón-Mateo C (2006) Causal agent of sharka disease: *Plum pox virus* genome and function of gene products. *EPPO Bulletin* 36: 229-238.
4. Malpica JM, Fraile A, Moreno I, Obies CI, Drake JW, et al. (2002) The rate and character of spontaneous mutation in an RNA virus. *Genetics* 162: 1505-1511.
5. Domingo E, Holland JJ (1997) RNA virus mutations and fitness for survival. *Annual Review of Microbiology* 51: 151-178.
6. Eigen M (1996) On the nature of virus quasispecies. *Trends in Microbiology* 4: 216-218.
7. Biebricher CK, Eigen M (2006) What is a quasispecies? *Current Topics in Microbiology and Immunology* 299: 1-31.
8. Holmes EC, Moya A (2002) Is the quasispecies concept relevant to RNA viruses? *Journal of Virology* 76: 460-465.
9. Jenkins GM, Worobey M, Woelk CH, Holmes EC (2001) Evidence for the non-quasispecies evolution of RNA viruses. *Molecular Biology and Evolution* 18: 987-994.
10. Sáenz P, Quiot L, Quiot J-B, Candresse T, García JA (2001) Pathogenicity determinants in the complex virus population of a *Plum pox virus* isolate. *Molecular Plant-Microbe Interactions* 14: 278-287.
11. Sáenz P, Cervera MT, Dallot S, Quiot L, Quiot JB, et al. (2000) Identification of a pathogenicity determinant of *Plum pox virus* in the sequence encoding the C-terminal region of protein P3+6K1. *Journal of General Virology* 81: 557-566.
12. López-Moya JJ, García JA (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 Research* 68: 99-107.
13. Rohozkova J, Navratil M (2011) P1 peptidase--a mysterious protein of family Potyviridae. *Journal of Biosciences* 36: 189-200.
14. Verchot J, Carrington JC (1995) Debilitation of plant potyvirus infectivity by P1 proteinase-inactivating mutations and restoration by second-site modifications. *Journal of Virology* 69: 1582-1590.
15. Verchot J, Koonin EV, Carrington JC (1991) The 35-kDa protein from the N-terminus of a potyviral polyprotein functions as a third virus-encoded proteinase. *Virology* 185: 527-535.
16. Brantley JD, Hunt AG (1993) The N-terminal protein of the polyprotein encoded by the potyvirus tobacco vein mottling virus is an RNA-binding protein. *Journal of General Virology* 74: 1157-1162.
17. Soumounou Y, Laliberté J-F (1994) Nucleic acid-binding properties of the P1 protein of turnip mosaic potyvirus produced in *Escherichia coli*. *Journal of General Virology* 75: 2567-2573.
18. Verchot J, Carrington JC (1995) Evidence that the potyvirus P1 proteinase functions in trans as an accessory factor for genome amplification. *Journal of Virology* 69: 3668-3674.

19. Pruss G, Ge X, Shi XM, Carrington JC, Vance VB (1997) Plant viral synergism: The potyviral genome encodes a broad-range pathogenicity enhancer that transactivates replication of heterologous viruses. *Plant Cell* 9: 859-868.
20. Rajamäki ML, Kelloniemi J, Alminaitte A, Kekarainen T, Rabenstein F, et al. (2005) A novel insertion site inside the potyvirus P1 cistron allows expression of heterologous proteins and suggests some P1 functions. *Virology* 342: 88-101.
21. Valli A, Martín-Hernández AM, López-Moya JJ, García JA (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. *Journal of Virology* 80: 10055-10063.
22. Nakahara KS, Shimada R, Choi S-H, Yamamoto H, Shao J, et al. (2010) Involvement of the P1 cistron in overcoming eIF4E-mediated recessive resistance against *Clover yellow vein virus* in pea. *Molecular Plant-Microbe Interactions* 23: 1460-1469.
23. Adams MJ, Antoniw JF, Fauquet CM (2005) Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Archives of Virology* 150: 459-479.
24. Valli A, López-Moya JJ, García JA (2007) Recombination and gene duplication in the evolutionary diversification of P1 proteins in the family *Potyviridae*. *Journal of General Virology* 88: 1016-1028.
25. Salvador B, Delgadillo MO, Saénz P, García JA, Simón-Mateo C (2008) Identification of *Plum pox virus* pathogenicity determinants in herbaceous and woody hosts. *Molecular Plant-Microbe Interactions* 21: 20-29.
26. Chiang CH, Lee CY, Wang CH, Jan FJ, Lin SS, et al. (2007) Genetic analysis of an attenuated *Papaya ringspot virus* strain applied for cross-protection. *European Journal of Plant Pathology* 118: 333-348.
27. Salvador B, Saenz P, Yanguéz E, Quiot JB, Quiot L, et al. (2008) Host-specific effect of P1 exchange between two potyviruses. *Molecular Plant Pathology* 9: 147-155.
28. Jridi C, Martin JF, Marie-Jeanne V, Labonne G, Blanc S (2006) Distinct viral populations differentiate and evolve independently in a single perennial host plant. *Journal of Virology* 80: 2349-2357.
29. Vancanneyt G, Schmidt R, O'Connor-Sanchez A, Willmitzer L, 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. *Molecular and General Genetics* 220: 245-250.
30. Moxon S, Schwach F, MacLean D, Dalmay T, Studholme DJ, et al. (2008) A toolkit for analysing large-scale plant small RNA datasets. *Bioinformatics* 24: 2252-2253.
31. Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589-595.

Table 1 Genome localization of nucleotide and amino acid differences between different PPV-PS subisolates and the cDNA clone PSeg

	<u>5'NCR</u>	<u>P1</u>				<u>HCPPro</u>				<u>CI</u>		<u>VPg</u>		<u>NIa-Pro</u>		<u>NIb</u>		<u>CP</u>				
Virus^a	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)
PS MCI	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^b	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)
1.3.1.	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)
2.1.1.	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)

Grey and white rows indicates virus causing severe and mild symptoms, respectively

^a Upper row indicates nucleotide position in the genome and, in parenthesis, the amino acid position in each protein

^b PSeg also differed from all subisolates in three nucleotide change at positions 2921, 2927 and 2936 that do not alter the amino acid sequence, which were introduced during the construction of the cDNA clone to facilitate the cloning of an intron [12], and in an extra T in the 3'NCR, at position 9629, which was introduced accidentally during the cloning procedure

Figure Legends:

Figure 1. Infectivity of PPV-PS mutants in *N. clevelandii*. (A) Photos of systemically infected leaves of plants inoculated with the indicated mutants or the subisolate PPV-PS 1.3.1 taken at 21 d.p.i. are shown in the panel. The severity of the symptoms is ranked from barely detectable (+/-) to the most intense chlorotic mottling (+++). (B and C) Virus accumulation in young systemically infected leaves of *N. clevelandii* biolistically- (B) or manually- (C) inoculated was determined by DASi-ELISA and western blot at 21 d.p.i. In (B), 1.3.1 subisolate was manually inoculated since 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.

Figure 2. Infectivity of PPV-PS mutants in *N. occidentalis*. (A) Photos of inoculated leaves taken at 10 d.p.i. are shown in the panel. Symptoms induced by the different viruses are described below the photos. 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 d.p.i.) determined by western blot.

Figure 3. Infectivity of PPV-PS mutants in peach GF305 seedlings. (A) Photos of systemically infected leaves taken at 35 d.p.i. 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 biolistically inoculated peach GF305 seedlings was determined by DASi-ELISA and western blot at 35 d.p.i. 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.

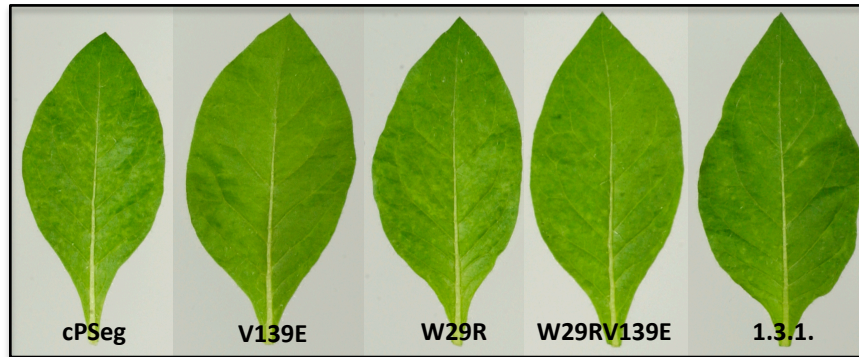
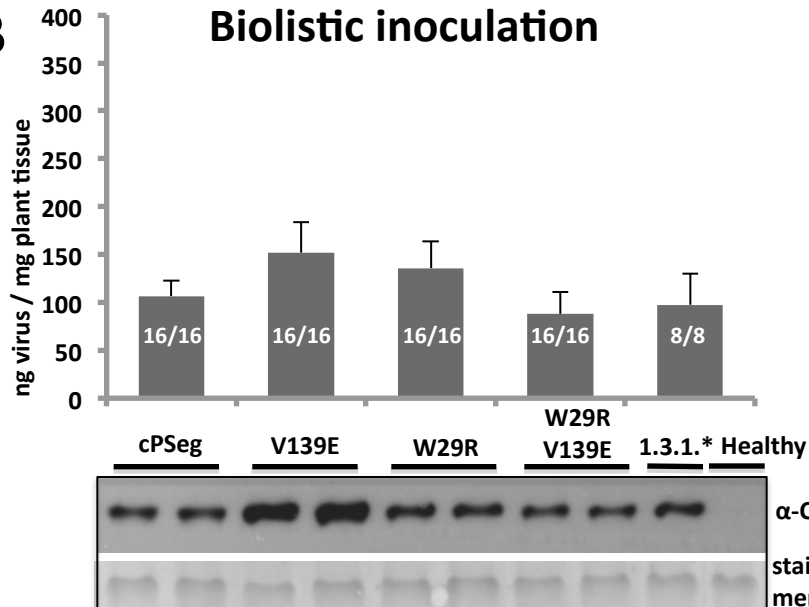
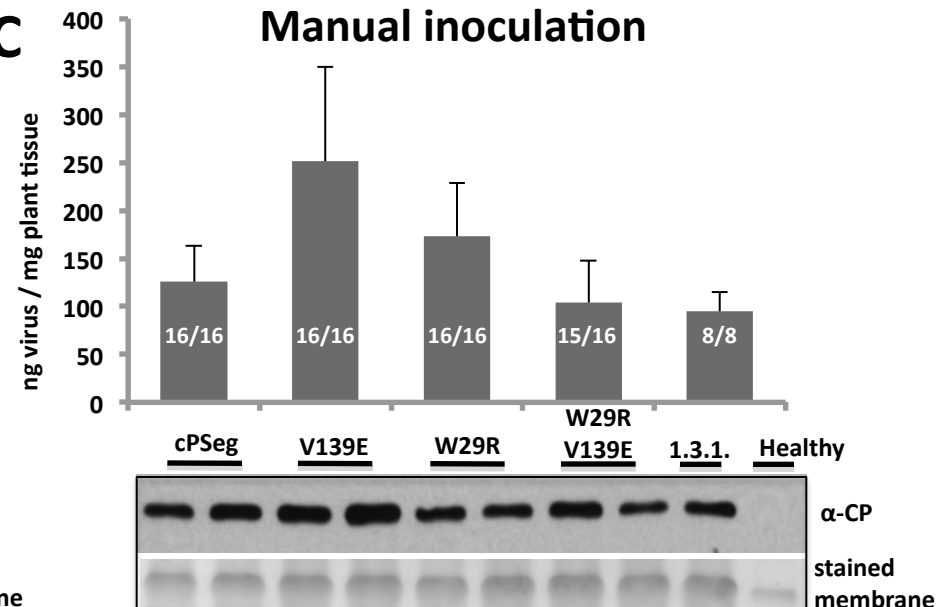
Figure 4. Heterogeneity map of the PPV-PS population. Height of the lines represents the heterogeneity values (H_i) of the indicated position for siRNAs that match the viral plus (up lines) or minus (down lines) strand. Only values of positions covered by more than 40 siRNA reads and more than 3% of 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 differentiated PS107 from the rest of subisolates (green lines). Black lines correspond to heterogeneities that did not segregate in PS subisolates. Labels show the two most abundant nucleotides at the indicated position.

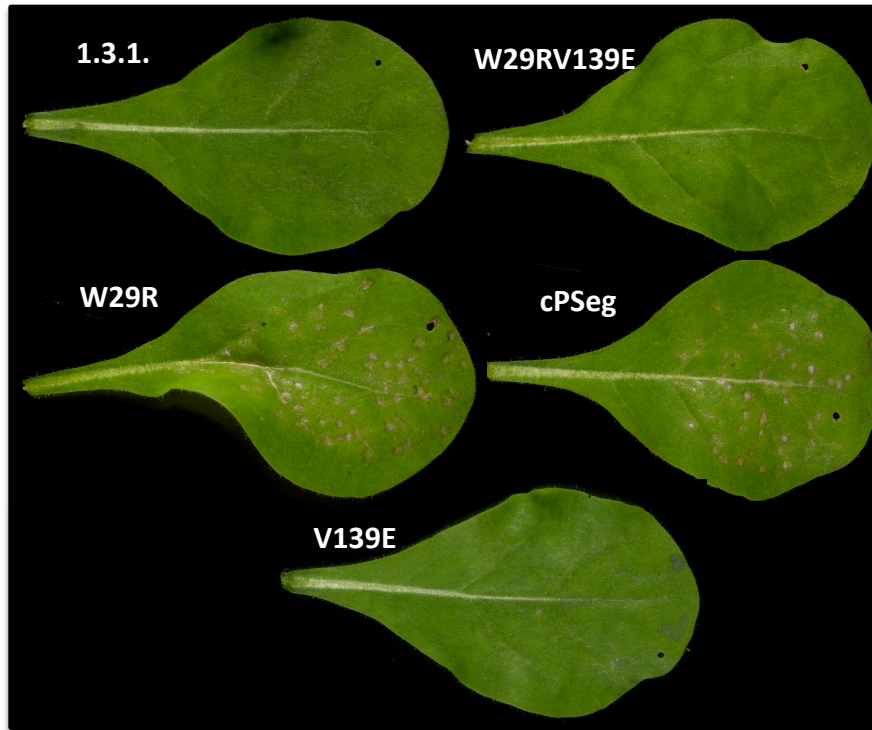
Supplementary Figure 1. Multiple amino acid alignment of the P1 proteins of 97 isolates belonging to 7 *Plum pox virus* strains recovered from public databases. The amino acid conservation at positions W29 and V139E is highlighted.

Supplementary Figure 2. Alignment of small RNAs to 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

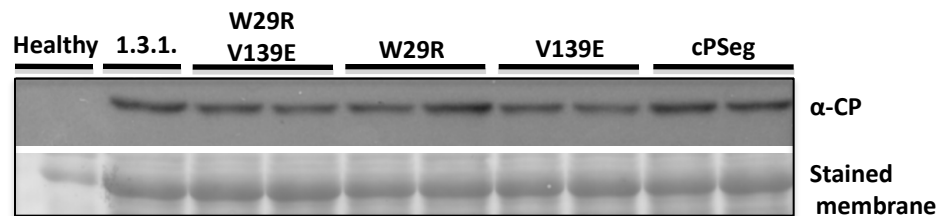
A

Virus	Symptoms (systemically infected leaves)	
1.3.1	Very mild chlorotic mottling	+
cPSeg	Strong chlorotic mottling	+++
V139E	Very mild chlorotic mottling	+/-
W29R	Strong chlorotic mottling	+++
W29RV139E	Mild chlorotic mottling	+

**B****C**

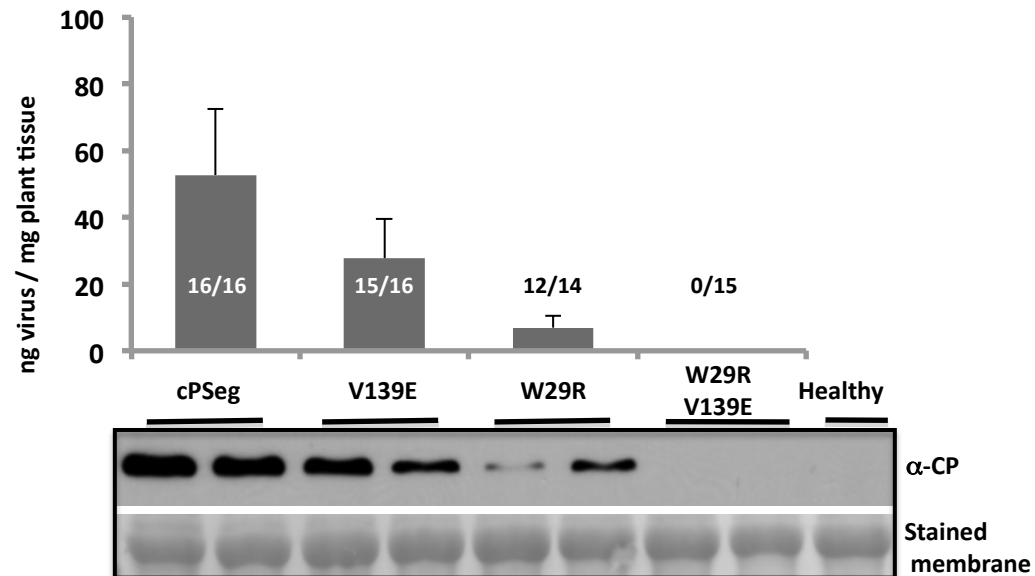
A

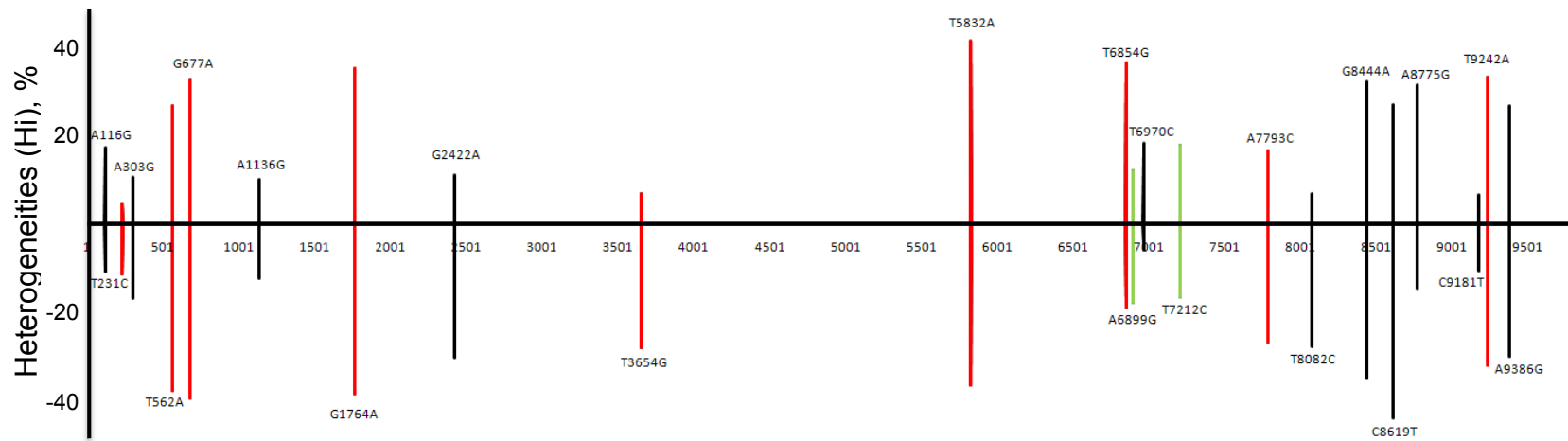
Virus	Lesions (Inoculated leaves)
1.3.1.	-
cPSeg	++
V139E	-
W29R	++
W29RV139E	-

B

A

Virus	Symptoms (systemically infected leaves)	
1.3.1.	-	-
cPSeg	Strong vein clearing and leaf distortion	+++
V139E	Mild vein clearing and leaf distortion	++
W29R	Very mild vein clearing and leaf distortion	+
W29RV139E	-	-

**B**



10 20 30 40 50 60 70 80 90 100
cP9aa
809_US-1020pe_EF611244_D
817_HU-1013p1_M92280_M
684_EG-1001ap_D0431465_E
685_US-1007pe_D0465242_D
831_UN-1006un_X81083_D
686_US-1008pe_D0465243_D
833_IT-1010sw_Y09851_C
852_US-1021pe_EF611246_D
853_US-1022pe_EF611247_D
854_US-1023pe_EF611248_D
856_FR-1018ap_X16415_D
559_SK-1002ap_AY028309_R
807_US-1018pe_EF611242_D
555_TR-1001ap_AY677114_T
718_EG-1002ap_AM157175_E
146_EG-300p1_G0461889_Re
245_IT-19pe_M_Pull1
580_US-1001p1_Af401296_D
603_GR-1003un_PM955843_M
753_US-1009pe_EF640933_D
754_US-1010pe_EF640934_D
755_US-1011pe_EF640935_D
756_US-1012pe_EF640936_D
757_US-1013pe_EF640937_D
758_US-1014pe_EF640938_D
759_US-1015pe_EF640939_D
771_TR-1005ap_EU734799_T
760_PL-1001p1_EU117116_R
773_TR-1006p1_EU734799_T
776_TR-1007p1_EU734802_T
768_TR-1003ap_EU734794_T
626_US-1006pe_Af401295_D
648_CA-1000p1_AY912055_M
649_CA-1001pe_AY912056_D
650_CA-1002pe_AY912057_D
651_CA-1003pe_AY912058_D
652_CA-1004p1_AY953261_D
653_CA-1005pe_AY953262_D
654_CA-1006pe_AY953263_D
655_CA-1007pe_AY953264_D
656_CA-1008pe_AY953265_D
657_CA-1009p1_AY953266_D
802_RS-1014un_A3243957_M
658_CA-1010pe_AY953267_D
803_MD-1000un_AY184478_C
805_UN-1000un_D13751_D
806_US-1017pe_EF611243_D
808_US-1019pe_EF611244_D
787_IT-1007ap_G0550510_D
815_HU-1012a1_G0550513_D
789_IT-1008p1_G0550512_D
790_IT-1009ap_G0550514_R
786_GR-1005pe_G0550509_M
527_IT-1001pe_AY917135_M
788_BG-1013p1_G0550511_M
791_EG-1003ap_G0550515_E
526_IT-1000sw_G0565151_C
956_SK-23pe_D_Pull1
147_SK-300p1_G0461890_D
148_RS-300pe_G0474956_Re

110 120 130 140 150 160 170 180 190 200
cP9aa
809_US-1020pe_EF611244_D
817_HU-1013p1_M92280_M
684_EG-1001ap_D0431465_E
685_US-1007pe_D0465242_D
831_UN-1006un_X81083_D
686_US-1008pe_D0465243_D
833_IT-1010sw_Y09851_C
852_US-1021pe_EF611246_D
853_US-1022pe_EF611247_D
854_US-1023pe_EF611248_D
856_FR-1018ap_X16415_D
559_SK-1002ap_AY028309_R
807_US-1018pe_EF611242_D
555_TR-1001ap_AY677114_T
718_EG-1002ap_AM157175_E
146_EG-300p1_G0461889_Re
245_IT-19pe_M_Pull1
580_US-1001p1_Af401296_D
603_GR-1003un_PM955843_M
753_US-1009pe_EF640933_D
754_US-1010pe_EF640934_D
755_US-1011pe_EF640935_D
756_US-1012pe_EF640936_D
757_US-1013pe_EF640937_D
758_US-1014pe_EF640938_D
759_US-1015pe_EF640939_D
771_TR-1005ap_EU734799_T
760_PL-1001p1_EU117116_R
773_TR-1006p1_EU734799_T
776_TR-1007p1_EU734802_T
768_TR-1003ap_EU734794_T
626_US-1006pe_Af401295_D
648_CA-1000p1_AY912055_M
649_CA-1001pe_AY912056_D
650_CA-1002pe_AY912057_D
651_CA-1003pe_AY912058_D
652_CA-1004p1_AY953261_D
653_CA-1005pe_AY953262_D
654_CA-1006pe_AY953263_D
655_CA-1007pe_AY953264_D
656_CA-1008pe_AY953265_D
657_CA-1009p1_AY953266_D
802_RS-1014un_A3243957_M
658_CA-1010pe_AY953267_D
803_MD-1000un_AY184478_C
805_UN-1000un_D13751_D
806_US-1017pe_EF611243_D
808_US-1019pe_EF611244_D
787_IT-1007ap_G0550510_D
815_HU-1012a1_G0550513_D
789_IT-1008p1_G0550512_D
790_IT-1009ap_G0550514_R
786_GR-1005pe_G0550509_M
527_IT-1001pe_AY917135_M
788_BG-1013p1_G0550511_M
791_EG-1003ap_G0550515_E
526_IT-1000sw_G0565151_C
956_SK-23pe_D_Pull1
147_SK-300p1_G0461890_D
148_RS-300pe_G0474956_Re

