Geographically and temporally distant natural recombinant isolates of *Plum pox virus* (PPV) are genetically very similar and form a unique PPV subgroup

Miroslav Glasa,1 Laszlo Palkovics,2 Petr Kominek,3 Gérard Labonne,4 Soňa Pittnerová,1 Otakar Kúdela,1 Thierry Candresse5 and Zdeno Šubr1

1Institute of Virology, Department of Plant Virology, Slovak Academy of Sciences, Dúbravská cesta 9, 84505 Bratislava, Slovakia
2Agricultural Biotechnology Center, Environmental Biosafety Institute, PO Box 411, H-2101 Godollo, Hungary
3Research Institute of Crop Production, Drnovská 507, 161 06 Prague – Ruzyně, Czech Republic
4Equipe Épidémiologie – Virus/Vecteurs, UMR BGPI CIRAD TA 41/K, Campus International de Baillarguet, 34398 Montpellier, France
5UMR GDPP, INRA et Université Bordeaux 2, IBVM, Centre INRA de Bordeaux, BP 81, 33883 Villenave d’Ornon Cedex, France

Natural recombinant *Plum pox virus* (PPV) isolates were detected in Albania, Bulgaria, Czech Republic, Germany, Hungary and Slovakia. Despite different geographical origins and dates of isolation, all the recombinant isolates were closely related at the molecular level and shared the same recombination breakpoint as well as a typical signature in their N-terminal coat protein sequence, suggesting a common origin. Biological assays with four recombinant isolates demonstrated their capacity to be aphid-transmitted to various *Prunus* hosts. One of these isolates had a threonine-to-isoleucine mutation in the conserved PTK motif of its HC-Pro and showed a drastically decreased, although not abolished, aphid transmissibility. The complete genome sequence of one of the recombinant isolates, BOR-3, was determined, as well as some partial sequences in the HC-Pro and P3 genes for additional natural recombinant isolates. Analysis of the phylogenetic relationships between the recombinant isolates and other sequenced PPV isolates confirmed that the recombinant isolates form a phylogenetically homogeneous lineage. In addition, this analysis revealed an ancient recombination event between the PPV-D and M subgroups, with a recombination breakpoint located in the P3 gene. Taken together, these results indicate that recombinant isolates represent an evolutionarily successful, homogeneous group of isolates with a common history and unique founding recombination event. The name PPV-Rec is proposed for this coherent ensemble of isolates.

INTRODUCTION

Due to the absence of proofreading activity in their RNA replicase, plant RNA viruses have a high genetic variation potential (Garcia-Arenal et al., 2001, 2003). In addition to mutation, recombination is generally thought to be an important source of variability and an efficient tool to repair viral RNA genomes. Recombination events have undoubtedly contributed to the evolution of several plant viruses, and may have played an important role in virus adaptation (Simon & Bujarski, 1994; Lai, 1995; Nagy & Simon, 1997; Worobey & Holmes, 1999).

RNA recombination is a process in which two non-contiguous RNA segments are joined. Both homologous and non-homologous recombination can occur (Lai, 1992). RNA template-switching is thought to represent the most common recombination mechanism in RNA viruses. During replication, the virus replicase ceases elongation of the nascent strand, dissociates from the template strand,
switches to another template or to another site on the same template while still associated to the nascent strand and finally resumes elongation (Simon & Bujarski, 1994; Nagy & Simon, 1997; Aaziz & Tepefer, 1999).

Despite mounting evidence for the existence of recombination events in the evolutionary history of many RNA viruses, it is apparent that recombination rates, or at least the frequencies of recombinant isolates in virus populations, vary extensively. Reports on natural recombinant isolates within the genus Potyvirus are still limited to a few species (Cervera et al., 1993; Revers et al., 1996; Bousalem et al., 2000, 2003; Glais et al., 2002; Glasa et al., 2002b; Tomimura et al., 2003; Moreno et al., 2004; Krause-Sakate et al., 2004; Desbiez & Lecoq, 2004).

**Plum pox virus** (PPV) is the causal agent of the economically important sharka disease of stone fruit trees in most European countries and, sporadically, outside Europe. The virus is efficiently transmitted by many aphid species (Homoptera: Aphidae) in a non-persistent manner (Labonne et al., 1995). PPV is a member of the genus Potyvirus in the family Potyviridae. Potyviruses have a single-stranded plus-sense RNA genome of about 10 kb harbouring a single open reading frame (ORF). Two main subgroups of PPV isolates have been recognized, PPV-M and PPV-D (Candresse et al., 1998). Two additional minor subgroups are represented by the geographically limited El Amar isolates from Egypt (PPV-EA; Wetzel et al., 1991) and by cherry-adapted isolates (PPV-C; Nemchinov et al., 1998). An unusual PPV isolate reported from Canada (James et al., 2003) may represent a distinct subgroup.

A large number of partial PPV sequences are available, but these are mostly limited to partial or complete capsid protein (CP) gene sequences and to the adjacent 3’ non-coding region (NCR). Consequently, until recently the techniques used for PPV subgroup discrimination focused exclusively on this short region of the genome (Wetzel et al., 1992; Bousalem et al., 1994; Candresse et al., 1998). More recently, the P3-6K1 region has been the target of variability studies, and typing methods for the P3-6K1 and CI regions have been developed, allowing PPV subgroup discrimination independently of the CP (Glasa et al., 2002a, b).

In the present study, the frequent occurrence of closely related recombinant PPV isolates in several European countries is demonstrated and the complete genome sequence of a representative recombinant isolate, named BOR-3, determined. The results demonstrate that these recombinant isolates form a coherent and evolutionarily linked group for which the name PPV-Rec is proposed.

**METHODS**

**Virus isolates.** Natural field PPV isolates were collected from different Prunus sources (plum, blackthorn, peach, apricot) from 1993 to 2003 at a number of distinct geographical locations throughout Europe (Table 1).

**Typing of PPV isolates.** PPV subgroup typing was performed in three genomic regions (CP, CI and P3-6K1) of each PPV isolate by RT-PCR followed by RFLP analysis, as described previously (Glasa et al., 2002b). In the case of the CP region, serological typing was also performed using monoclonal antibodies (mAbs) specific to the PPV-M or the D subgroup (Candresse et al., 1998) in triple-antibody-sandwich (TAS)-ELISA assays, as recommended by the suppliers.

**Determination of the complete genome sequence of the BOR-3 isolate.** Total RNAs were extracted from lyophilized *Nicotiana benthamiana* leaves infected with the BOR-3 isolate using the RNeasy Plant Mini kit (Qiagen). Oligo(dT) or pdN6 random-primed cDNA synthesis was carried out using avian myeloblastosis virus (AMV) reverse transcriptase (Promega). Seven pairs of primers were designed based on available PPV sequences in order to amplify seven overlapping PCR fragments covering the whole genome (details are given in a Supplementary Table available in JGV Online). All PCR amplifications were performed using the proofreading TaKaRa Ex Taq polymerase (Takara Bio). PCR products were then sequenced directly using the respective PCR primers or custom-made internal primer when required. Alternatively, PCR products were ligated into the pGEM-T Easy vector (Promega), transformed into *Escherichia coli* JM109 cells and recombinant plasmids were sequenced using a MegaBACE 1000 DNA Analysis system (Amersham Biosciences). The complete BOR-3 genom sequence was finally assembled from the overlapping cDNA sequences. The organization of the BOR-3 genome was inferred by analogy with those of other fully sequenced PPV isolates: PS (AJ243957), SK68 (M92280), Dideron (X16415), NAT (NC_001445), SC (X81083), PENN-1 (AF401295), PENN-2 (AF401296), SoC (AY184478) and SwC (Y09851).

**Partial sequencing of PPV isolates and phylogenetic analyses.** To amplify various genomic regions, a two-step RT-PCR protocol was used. The first-strand cDNA was synthesized by reverse transcription of total RNA using pdN6 random primers. All individual PCRs were then performed using the same RT preparation. PCR was performed using the TaKaRa Ex Taq polymerase. For several PPV isolates, the cDNA fragment spanning the (Cter)NIb-(Nter)CP region corresponding to nucleotides 8339–8912 (numbered according to PPV-BOR-3, AJ028309) was amplified and sequenced using the PCR primer P4 (5’-TGCCTTTACCAAGCTGGCACTG-3’) in combination with either RD5 (5’-GGAAGCGTGGGGTATAAGGAG-3’) or RM5 (5’-GAGGGCCGTGGGCTACAAAG-3’, PPV-D specific) or RM5 (5’-GAGGGCTGGGGCTACAAAG-3’, PPV-M specific). The following conditions (Technie Genius; Merck) were used: initial denaturation at 94 °C for 5 min, 35 cycles of 94 °C for 45 s, 54 °C for 45 s and 72 °C for 1 min and a final extension at 72 °C for 10 min.

In addition, sequences in the HC-Pro and P3 genomic regions of a few isolates were also determined on RT-PCR-amplified cDNAs. Primers HC-RC (5’-GGATCAGTCTCAACTGGAATGC-3’) and P3-RC (5’-CTGCGATTCTCAAGATGTCAGAG-3’) were used to amplify a (Cter)HC-(Nter)P3 fragment corresponding to nt 2337–2390 under the same PCR conditions as for (Cter)NIB-(Nter)CP.

Primers 5-HC (5’-CTTGTTGACCTACGGTTAAGG-3’) and 3-HC (5’-GCTGTTGGGCAAACTCCACATAC-3’) were designated for amplification of two-thirds of the HC gene (nt 1014–2062). The PCR targeting this region was carried out under the following cycling conditions: denaturation at 94 °C for 5 min followed by 35 cycles (94 °C for 1 min, 53 °C for 45 s, 72 °C for 75 s). The final extension step at 72 °C lasted 15 min.

The PCR products were gel-purified (QIAprep Spin Miniprep kit; Qiagen) and sequenced directly. When necessary, PCR products were inserted into the cloning vector pGEM-T Easy. Partial sequences
were compared to sequences available in GenBank (http://www.ncbi.nlm.nih.gov). Sequence analyses and comparisons were performed using the CLUSTAL X program (Thompson *et al.*, 1997) or the GAP program of the Genetics Computer Group sequence analysis software version 10.0 (Devereux *et al.*, 1984). The alignments were used as input data to construct phylogenetic trees with the neighbour-joining

### Table 1. Typing of PPV field isolates using assays targeting the P3-6K1, CI and CP genomic regions

Typing tests were performed in most cases using fresh or lyophilized infected leaf samples from the original *Prunus* hosts in order to avoid the accumulation of mutations during transfer to, and propagation in, experimental herbaceous hosts. In a few cases, isolates were mechanically transmitted to *Nicotiana benthamiana* prior to testing. TAS-ELISA and Cter CP typing techniques were described by Candresse *et al.* (1998); CI and P3-6K1 were described by Glasa *et al.* (2002b). NT, Not tested.

<table>
<thead>
<tr>
<th>Isolate(s)</th>
<th>Origin</th>
<th>Original host, year of isolation</th>
<th>Type</th>
<th>Subgroup</th>
<th>Accession number(s)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BOR-3</td>
<td>Slovakia</td>
<td>Apricot, 1996</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY028309a</td>
</tr>
<tr>
<td>VAR-2</td>
<td>Slovakia</td>
<td>Peach, 1996</td>
<td>M M M M</td>
<td>M</td>
<td>AY324837d</td>
</tr>
<tr>
<td>CAH-2</td>
<td>Slovakia</td>
<td>Apricot, 1996</td>
<td>M M M M</td>
<td>M</td>
<td>AY324838c</td>
</tr>
<tr>
<td>DPLA</td>
<td>Slovakia</td>
<td>Peach, 2002</td>
<td>M M M M</td>
<td>M</td>
<td>AY324840d</td>
</tr>
<tr>
<td>Kr-4</td>
<td>Slovakia</td>
<td>Apricot, 1995</td>
<td>M M M M</td>
<td>M</td>
<td>AY324841c</td>
</tr>
<tr>
<td>BIII/2</td>
<td>Slovakia</td>
<td>Plum, 1996</td>
<td>D D D D</td>
<td>D</td>
<td>AY553370c</td>
</tr>
<tr>
<td>PEZ-1</td>
<td>Slovakia</td>
<td>Plum, 2002</td>
<td>D D D D</td>
<td>D</td>
<td>AY553371c</td>
</tr>
<tr>
<td>Vis-2</td>
<td>Slovakia</td>
<td>Apricot, 2002</td>
<td>D D D D</td>
<td>D</td>
<td>AY324844d</td>
</tr>
<tr>
<td>Slivoň</td>
<td>Czech Rep.</td>
<td>Plum, 1995</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY324843c</td>
</tr>
<tr>
<td>Horomerice</td>
<td>Czech Rep.</td>
<td>Plum, 2002</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY324844d, AY553377d</td>
</tr>
<tr>
<td>302</td>
<td>Czech Rep.</td>
<td>Plum, 1997</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY324845e</td>
</tr>
<tr>
<td>Nectagrand</td>
<td>Czech Rep.</td>
<td>Peach, 1997</td>
<td>M M M M</td>
<td>M</td>
<td>AY324842c</td>
</tr>
<tr>
<td>Spy4</td>
<td>Czech Rep.</td>
<td>Plum, 2002</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY324847c</td>
</tr>
<tr>
<td>Spy1, Spy2, Spy3, Olo1, Olo2, Olo3, HoK1, HoK2</td>
<td>Czech Rep.</td>
<td>Plum, 2002</td>
<td>D D D D</td>
<td>D</td>
<td>AY324848d, AY553378d</td>
</tr>
<tr>
<td>BULG</td>
<td>Bulgaria</td>
<td>Plum, 2001</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY324846c, AY553376d</td>
</tr>
<tr>
<td>89-006</td>
<td>Bulgaria</td>
<td>Plum, 1989</td>
<td>NT M M M M</td>
<td>M</td>
<td>AY324842c</td>
</tr>
<tr>
<td>ASL-1</td>
<td>Germany</td>
<td>Plum, 2002</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY324847c</td>
</tr>
<tr>
<td>143A</td>
<td>Germany</td>
<td>Plum, 1995</td>
<td>M M D D</td>
<td>Rec</td>
<td>AY324848d, AY553375d, AY553372d</td>
</tr>
<tr>
<td>ASL-2, ASL-3, ASL-4, Jena, Mr1, Qlb</td>
<td>Germany</td>
<td>Plum, 2002</td>
<td>D D D D</td>
<td>D</td>
<td>AY553368e</td>
</tr>
<tr>
<td>ALB7</td>
<td>Albania</td>
<td>Plum, 2003</td>
<td>M M NT D D</td>
<td>D</td>
<td>AY553369c</td>
</tr>
<tr>
<td>ALB9</td>
<td>Albania</td>
<td>Plum, 2003</td>
<td>M M NT D D</td>
<td>Rec</td>
<td>AY553368e</td>
</tr>
<tr>
<td>ALB1, ALB2, ALB4, ALB5, ALB6, ALB8, ALB10</td>
<td>Albania</td>
<td>Plum, 2003</td>
<td>M M NT M M</td>
<td>M</td>
<td>AY553368e</td>
</tr>
<tr>
<td>ALB3</td>
<td>Albania</td>
<td>Plum, 2003</td>
<td>M M NT D D</td>
<td>Rec</td>
<td>AY553368e</td>
</tr>
<tr>
<td>Bt-H2</td>
<td>Hungary</td>
<td>Blackthorn, 1998</td>
<td>NT M NT D D</td>
<td>Rec</td>
<td>AJ566346b</td>
</tr>
<tr>
<td>Pd4</td>
<td>Hungary</td>
<td>Plum, 1996</td>
<td>NT M D D D</td>
<td>Rec</td>
<td>AJ566344b, AJ620687d</td>
</tr>
<tr>
<td>Pd31</td>
<td>Hungary</td>
<td>Plum, 1996</td>
<td>NT M D D D</td>
<td>Rec</td>
<td>AJ566345b, AJ620686d</td>
</tr>
<tr>
<td>Gr1, Gr2, Gr3, Gr4, Gr5</td>
<td>Greece</td>
<td>Peach, 2002</td>
<td>M M NT M M</td>
<td>M</td>
<td>AY553368e</td>
</tr>
</tbody>
</table>

*Accession numbers refer to sequences of the following regions: a, complete genomic sequence; b, (Cter)NIb–CP–3' UTR; c, (Cter)NIb–(Nter)CP; d, (Cter)HC–(Nter)P3; e, HC gene. Cter, C terminus; Nter, N terminus.*
distance method implemented in CLUSTAL X. The trees were visualized using the program TREEVIEW version 1.6.1. The recombinant sequences and the location of recombination breakpoints were detected using PHYLPRO version 1.0 (Weiller, 1998). Possible recombination events in the PPV genomes were also studied using the GENECONV version 1.81 software (http://www.math.wustl.edu/~sawyer/geneconv/) (Sawyer, 1989). MEGA2 (Kumar et al., 2001) was used to estimate mean intra- and intergroup nucleotide divergence values.

**Aphid transmission.** Experimental aphid transmission was performed under controlled conditions using a technique involving a controlled acquisition access period (Labonne et al., 1995) and laboratory-maintained clones of *Myzus persicae* Sulzer. Leaves from *N. benthamiana* infected with the respective PPV isolates were used as inoculum sources for aphid feeding. Four recombinant isolates were compared (BOR-3, BULG, Horomerice and 143A) and a French isolate 91.003, belonging to the PPV-M subgroup, was used as a control. Three-week-old seedlings of various *Prunus* genotypes, peach (*Prunus persica* cv. GF305), apricot (*Prunus armeniaca* cv. Manicot) and plum (*Prunus domestica* cv. Jullor), were used as test plants. Fifty aphids were deposited on each test plant after a 3 h starvation period followed by a 6 min acquisition period on the source leaves. Aphids were left overnight on the plants and then killed with an aphicide treatment. PPV was finally detected in the test plants by DAS-ELISA after a 3 week incubation period.

**Nucleotide sequence accession numbers.** The nucleotide sequences reported in this paper have been deposited in the GenBank database under accession numbers AY028309, AY553368–AY553377, AY324837–AY324848, AJ566344–AJ566346, AJ620686–AJ620687.

**RESULTS**

**Presence of natural recombinant PPV isolates in several European countries**

The frequent occurrence of natural recombinant PPV isolates has recently been detected in Slovakia (Glasa et al., 2002b). In order to determine whether similar or other recombinant isolates are established in other European countries, PPV-infected samples were collected from a number of European countries (Albania, Bulgaria, the Czech Republic, Germany, Greece, Hungary) and submitted to typing assays targeting different parts of the PPV genome. Of the 65 field samples analysed, 23 presented discrepancies between the results derived from CP-based typing assays (ELISA with subgroup-specific mAbs, RT-PCR of the 3′ end of the CP followed by *Rsa*I digestion; Candresse et al., 1998) and typing assays targeting upstream genomic regions: the RFLP analysis of the P3-6K1 region and subgroup-specific RT-PCR in the *CI* gene (Glasa et al., 2002b) (Table 1). On the other hand, no discrepancies were observed between the two techniques targeting the CP gene region or between the two techniques targeting the P3-6K1 and CI genomic regions (Table 1).

In previous studies, a recombination breakpoint was detected in the C-terminal part of the *NIb* gene in the o6 isolate from former Yugoslavia (Cervera et al., 1993) and in Slovak recombinant isolates (Glasa et al., 2002b). Therefore, a primer pair targeting the (Cter)NIb-(Nter)CP region was used to amplify cDNAs from 11 of the tentative recombinant isolates. The resulting RT-PCR products were sequenced to assess the genetic variability of these isolates and to determine whether, as for the o6 and recombinant Slovak isolates, these isolates have a recombination breakpoint in this region. In addition, six PPV-M and two PPV-D isolates from Bulgaria, the Czech Republic and Slovakia (Table 1) were also sequenced in this region as controls, extending the number of non-recombinant sequences available for the subsequent phylogenetic analyses.

Analysis of the last 238 nt of the *NIb* gene and the first 336 nt of the *CP* gene demonstrated that, despite different geographical origins and date of isolation, all tentative recombinants sequenced showed a similar recombination breakpoint situated in the C terminus of *NIb*. This breakpoint is in the same position as the one previously described for the o6 and recombinant Slovak isolates (Supplementary Fig. A, available in JGV Online). Thus, despite testing a relatively small number of isolates, the further presence of recombinant isolates was demonstrated in Albania, Bulgaria, Czech Republic, Germany and Hungary (Table 1).

In addition to sharing the same recombination breakpoint, all the recombinant isolates sequenced shared a common typical signature in their N-terminal CP sequence. Five amino acids (*K*2814, *I*2848, *T*2852, *I*2868, *T*2878) were specifically conserved in all the recombinant sequences and were not found in the sequence of the reference PPV-M and D isolates characterized to date. The single exception was the Pd4 isolate, which had an alanine instead of threonine at amino acid position 2878. Thus, these recombinant isolates appear to form a coherent ensemble of phylogenetically related PPV isolates. This conclusion is supported by phylogenetic reconstructions performed using the hypervariable N-terminal region of the CP gene (Fig. 1). Indeed, all the recombinant isolates cluster on a highly phylogenetically supported branch, clearly distinguished from the typical PPV-M isolates. On the other hand, similarly to what was observed in the PPV-D and M subgroups, no obvious historical or geographical structure of the recombinant isolates was observed (Fig. 1; results not shown).

It is noteworthy that, based on their CP sequences, several previously sequenced isolates, Bulgarian (X57976), CG (X81082), CGG-M2 (AY450594), CGG-M3 (AY450595) and CGG-M6 (AY450596), were originally described as belonging to the PPV-M subgroup cluster with the recombinants and are therefore very likely to be recombinant themselves (Fig. 1). It is, however, not possible to provide definite proof of this hypothesis, in that the available sequence information does not cover the *NIb* recombination breakpoint.

Mean nucleotide sequence divergence (diversity) levels within and between the three subgroups of PPV isolates (D, M, recombinants) were calculated using the MEGA2 program on alignments generated from the hypervariable N-terminal region of the CP gene (first 336 nt; Table 2). The largest intra-group diversity was observed within the
PPV-D subgroup, while the lowest diversity was observed for the recombinant isolate subgroup. In addition, the mean divergence between recombinants and PPV-M isolates was 6.4%, while the intra-subgroup divergence levels reached 1.5–2.4%. Both results are consistent with a relatively ancient recombination event, giving rise to the recombinant subgroup followed by its divergence. Based on the above results, all the recombinant isolates should therefore be classified as members of a unique PPV subgroup, for which the name PPV-Rec is proposed.

Table 2. Mean intra- and intersubgroup nucleotide sequence divergence levels in the N-terminal hypervariable region of the coat protein gene among available sequence of 64 PPV isolates, including isolates sequenced in this study

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>PPV-D</th>
<th>PPV-M</th>
<th>PPV-Rec</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPV-D (29 isolates)</td>
<td>0.024 ± 0.002</td>
<td>0.226 ± 0.019</td>
<td>0.219 ± 0.020</td>
</tr>
<tr>
<td>PPV-M (8 isolates)</td>
<td>0.019 ± 0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPV-Rec (27 isolates)</td>
<td></td>
<td></td>
<td>0.015 ± 0.003</td>
</tr>
</tbody>
</table>

Fig. 1. Phylogenetic tree of PPV isolates reconstructed using the nucleotide sequence encoding the hypervariable N-terminal region of the coat protein (first 336 nt). The sequences used were either recovered from the databases or determined in this work (Table 1). Bar, 0.1 substitutions per site. Bootstrap analysis with 1000 replicates was performed to assess the robustness of the branches. Only bootstrap values >70% are shown. The divergent PPV-SoC, SwC and El Amar were used as outgroup. Accession numbers of PPV isolates used (if not mentioned elsewhere) are SH (X81073), NL (X81074), NEB (X81075), OB1 (X81077), AL (X81078), SL (X81079), GSP (X81080), Li/H (X81081), DOH1 (X81084), PA (A000340), AT (X57975), o6 (S57404), El Amar (X56258), AIN (AF332871), Skierniewice (U27652), Chile112 (AF440741), Chile114 (AF440742), Chile116 (AF440743), Chile12 (AF440744), Chile20 (AF440745), Chile31 (AF440746), PEN Pch (AF354268), PEN Plm (AF354269), MNAT1 (AF360579), DCK1 (AF360580), BNE-10 (AF450311), LOZ-3 (AF450312), MYV-3 (AF450313), BRC-3 (AF421118), BRC-4 (AF421119), BRC-8 (AF421120), KN-1 (AF421121), KN-4 (AF421122) and KN-7 (AF421123).
Table 3. Transmission rates of recombinant and control PPV isolates on different Prunus species

Entries are number of infected plants/total number of plants. Different letters indicate significant differences (P < 0.05; exact test) between isolates for the same host plant.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>P. domestica</th>
<th>P. armeniaca</th>
<th>P. persica</th>
<th>Total Prunus</th>
<th>N. benthamiana</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Julius</td>
<td>Manicot</td>
<td>GF305</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BOR-3</td>
<td>9/20b</td>
<td>11/20b</td>
<td>1/20b</td>
<td>21/60 (35%)</td>
<td>6/10</td>
</tr>
<tr>
<td>BULG</td>
<td>8/20b</td>
<td>6/20ab</td>
<td>3/20b</td>
<td>17/60 (28%)</td>
<td>4/10</td>
</tr>
<tr>
<td>Horomerice</td>
<td>18/20c</td>
<td>19/20c</td>
<td>11/20b</td>
<td>48/60 (80%)</td>
<td>9/10</td>
</tr>
<tr>
<td>143A</td>
<td>0/20a</td>
<td>2/20a</td>
<td>1/20a</td>
<td>3/60 (5%)</td>
<td>2/10</td>
</tr>
<tr>
<td>91.003</td>
<td>16/20b</td>
<td>19/20c</td>
<td>7/20ab</td>
<td>42/60 (70%)</td>
<td>9/10</td>
</tr>
</tbody>
</table>

Biological properties of the recombinant isolates

Experimental aphid-transmission experiments were performed to evaluate the transmissibility of four recombinant isolates, as well as their ability to infect several Prunus species efficiently. The transmission experiments confirmed that all the recombinant PPV isolates tested were aphid-transmitted from infected N. benthamiana to the three Prunus genotypes tested. However, despite high similarity in the amino acid sequences of the N-terminal CP region and the presence of the conserved DAG motif (Atreya et al., 1995), different aphid-transmission rates were noted depending on the isolate (Table 3). The Horomerice isolate showed a particularly high efficiency of aphid transmission in comparison with the other isolates, while isolate 143A showed consistently lower transmission efficiency irrespective of the experimental host used. The observed variability in transmission efficiency cannot be attributed to differences in virus titre in the N. benthamiana leaves used as inoculum sources, because these showed similar virus concentrations as estimated by semi-quantitative DAS-ELISA assays (data not shown). In contrast, there were no differences in the symptoms caused by four recombinant isolates, or in the development and evolution of symptoms for the different inoculated Prunus species over a period of 1 year post-inoculation by aphid.

The nucleotide sequence of the HC-Pro gene of the four recombinant isolates tested was determined in order to evaluate their potential heterogeneity and to verify the presence of conserved amino acid motifs suggested to be involved in aphid transmission. Sequence comparison of the deduced sequences of the 326 N-terminal amino acids encoded by HC-Pro (three-quarters of HC-Pro including all the known motifs involved in aphid transmission) showed very high amino acid sequence identity levels between all four isolates (99.4–100%; nucleotide sequence identities 98.7–99.3%).

All conserved amino acid motifs of the potyvirus helper component were unaltered in the four PPV recombinants, with a single exception, a threonine-to-isoleucine mutation in the conserved Pro-Thr-Lys (PTK; Blanc et al., 1997) motif of isolate 143A as a consequence of an ACA to ATA codon change.

These experiments indicate that, in addition to individual differences, the recombinant isolates are transmissible by aphids to plum, apricot and peach species, and are thus able to spread in these hosts under natural epidemiological conditions.

Complete genomic sequence of the BOR-3 recombinant isolate

So far, nine complete sequences of PPV isolates belonging to the M, D and C subgroups are available. The complete genomic sequence of one of the recombinant isolates, BOR-3, was determined in order to compare it with the other fully sequenced isolates and to evaluate whether other recombination events have occurred in the evolutionary history of PPV isolates. The genomic RNA of BOR-3 (accession no. AY028309) is composed of 9786 nt, excluding the poly(A) tail, and has an organization similar to that reported for the other PPV isolates sequenced. The 5’ and 3’ NCR of BOR-3 are the same length as the previously characterized PPV isolates (146 and 217 nt long, respectively). The single large ORF of 9534 nt encodes a polyprotein of 3140 amino acids with a calculated molecular mass of 355 kDa. Nine potential cleavage sites for the viral proteases (Fig. 2) were identified at the expected positions in the predicted amino acid sequence of the BOR-3 polyprotein. The percentages of nucleotide sequence identity calculated for each BOR-3 gene against the conventional PPV-D, M and C isolates are presented in Fig. 2. All characteristic motifs of potyviral proteins fundamental for the virus cycle and for vector transmission are conserved in the BOR-3 genome (data not shown).

Comparison of the complete BOR-3 nucleotide sequence with other PPV isolates revealed 90.2–90.4% identity with isolates of the PPV-M subgroup (SK68, PS), 94.8–95.7% identity with isolates of the PPV-D subgroup (NAT, SC, Dideron, PENN-1, PENN-2) and only 78.1–78.2% identity with isolates of PPV-C subgroup (SoC, SwC). Analysis of the
phylogenetic relationship using whole-genome sequences confirmed that BOR-3 forms a phylogenetically distinct group with 100% bootstrap support (data not shown).

**Analysis of recombination events in complete PPV genomic sequences**

To identify other potential recombination events within the BOR-3 genome, all full-length PPV sequences, including BOR-3, were analysed using the PHYLPRO program. The phylogenetic correlation profiles are shown in Fig. 3. In the case of BOR-3, a clear peak indicated a recombination breakpoint around nt 8450, corresponding to the end of the Ni(b) gene and the previously identified recombination event (see above). Unexpectedly, a second putative recombination signal appeared in the 3′ part of the P3 gene, around nt 2813. A similarly positioned, but weaker, recombination signal was also observed for some of the other PPV isolates in the data set. As the detection of recombination events can be affected by the number of sequences in each phylogenetic cluster, additional PHYLPRO analyses were performed, keeping only a representative PPV-D and/or PPV-M sequence. Again, significant PHYLPRO recombination signals were observed in the same region for BOR-3 and for the other PPV subgroups (results not shown).

To attempt to validate this potential event, a second test for recombination was performed using Sawyer’s test as implemented in the GENECONV software (only PPV-M, D and BOR-3 isolates). The test was performed using a minimum mismatch penalty option on only the 5′ part of the PPV genome spanning nt 1–3939 and representing the 5′ NCR and the P1, HC-Pro, P3, 6K1 and (N-ter)CI genes. This test validated as highly significant the existence of a recombination event around position 2813 in BOR-3, but also in other PPV isolates representative of the PPV-D and M subgroups (probability for absence of recombination < 10^-4, result not shown).

Multiple alignments of the 5′ part of the PPV genome were analysed in more detail in order to understand the evolutionary history of this genomic region. A first observation revealed that the two regions upstream and downstream from the suspected recombination breakpoint around nt 2813 differ in the intensity of the phylogenetic signals they contain. The 5′ portion (nt 1–2813) contains only 63 phylogenetically informative sites (0.022 sites per nucleotide) while the 3′ portion (nt 2814–3939) contains 139 such sites (0.122 sites per nucleotide). Overall, this is a sixfold difference in density and indicates either different evolutionary histories or different selection pressures applied to these two regions.

**Fig. 2.** Polyprotein map of the PPV-BOR-3 isolate and percentages of nucleotide sequence identity between the BOR-3 genomic regions and those of other fully sequenced PPV isolates belonging to the PPV-D, PPV-M and PPV-C subgroups. Amino acid sequences of the putative cleavage sites are shown under the schematic representation of the genome. The length in amino acids of each functional product is indicated in the diagram. Nucleotide sequence identities between the functional products of PPV-BOR-3 and the corresponding regions of other fully sequenced PPV isolates from -M, -D and -C subgroups are shown in the table below the map.
The very different behaviour of these two regions was confirmed by the phylogenetic reconstructions presented in Fig. 4, in which the two trees are drawn to the same scale. The topologies of the two trees are essentially similar. The branches leading to the PPV-D and M subgroups are strongly supported by 99–100% bootstrap values. However, the trees are of distinctly different sizes (reflecting the difference in phylogenetic signal between the two regions identified above) due to variation in the length of the branch separating the PPV-M subgroup isolates PS and SK68.

These trends are confirmed by the calculations for the two regions of intra- and inter-subgroup diversity (mean nucleotide sequence divergence). Indeed, while the mean divergence between the PPV-D and M subgroup in the 1–2813 region was only 3.7%, it jumped to 13.8% in the 2814–3939 region (similar values for M/BOR-3 comparisons were 3.1 and 12.8%, respectively). Concurrently, the intra-subgroup diversities were similar in the two regions, with values of 1.6% (Table 4).

Taken together, the above results are consistent with the existence of an ancient recombination event with a breakpoint located in the P3 gene around position 2813, which resulted in an exchange of the 5’ part of the genome between the PPV-D and M subgroups (or between PPV-M and PPV-D). Such a hypothesis is consistent with the large differences in diversity and in phylogenetic signal between the two regions, as well as with the results of the two recombination analyses carried out with PHYLPilo and GENECONV.

To determine whether other recombinant isolates behave similarly to BOR-3 in the 5’ part of their genome, the genomic region spanning the C-terminal part of HC-Pro and the N-terminal part of P3 was amplified from five other recombinant isolates (Pd4, Pd31, BULG, Horomerice and 143A) and sequenced. All these isolates proved to be very closely related to BOR-3 in this region, with nucleotide sequence identity levels of 98–99.5%.

**DISCUSSION**

The accurate identification of specific isolate groups and the determination of their genetic variability are among the first steps in designing effective disease-control strategies. The frequent occurrence of recombinant PPV isolates has been overlooked for a long time, due to the use of typing methods focusing solely on the CP gene or protein. The approach supplementing conventional CP-based techniques with typing methods targeting other upstream genomic regions (Glasa et al., 2002a, b) has proven to be a powerful tool for PPV recombinant isolate identification.

The first reported PPV recombinants (Cervera et al., 1993; Glasa et al., 2001) were considered to be unusual isolates, not representative of PPV field populations. The screening of a large number of field samples presented here demonstrates that, on the contrary, recombinants are rather common, at least in central and southern Europe, representing about one-third of all isolates analysed.

Despite the fact that they were collected over a very large geographical area with different environments (central and south-east Europe) and over a relatively large time scale (1993–2003), the high genetic similarity found between all the recombinants analysed unambiguously supports the

---

**Table 4. Mean inter- and intrasubgroup diversity calculated for nt 1–2813 (top part of the matrix) and 2814–3939 (bottom part of the matrix)**

Values are means \(\pm\) SEM. For the intrasubgroup calculations, two values are given, the first value representing that calculated for nt 1–2813 and the second value (bold) that for nt 2814–3939. NA, Not applicable.

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>PPV-D</th>
<th>PPV-M</th>
<th>BOR-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPV-D</td>
<td>0.016 ± 0.001, 0.016 ± 0.002</td>
<td>0.037 ± 0.003</td>
<td>0.036 ± 0.003</td>
</tr>
<tr>
<td>PPV-M</td>
<td>0.138 ± 0.010</td>
<td>0.016 ± 0.003, 0.016 ± 0.004</td>
<td>0.031 ± 0.003</td>
</tr>
<tr>
<td>BOR-3</td>
<td>0.028 ± 0.003</td>
<td>0.128 ± 0.008</td>
<td>NA</td>
</tr>
</tbody>
</table>
hypothesis that they all share a common origin. This origin can be traced to a single recombination event between PPV-D and M subgroup isolates, with a breakpoint in the 3’ part of the Nlb gene between positions 8433 and 8449. This common evolutionary origin of all recombinant isolates analysed to date justifies the proposal that they be considered a distinct subgroup of PPV isolates, for which the name PPV-Rec is coined. It is always difficult to identify the direction of recombination events. The possibility remains that PPV-Rec isolates represent an ancestral group, while either PPV-D or PPV-M represents the recombinants deriving from the recombination event in the Nlb gene. The fact that PPV-Rec isolates have a lower genetic diversity (Table 2) suggests that they are younger, and therefore are the products of the recombination. However, confounding events, such as population bottlenecks occurring well after the recombination event (as a result of agricultural practices), might interfere with this tentative identification of the ‘younger’ and ‘older’ subgroups of isolates; therefore, the possibility will always remain that PPV-Rec represents an ancestral subgroup.

The current wide distribution of these recombinant isolates can be viewed as proof of their ecological success and the existence of a very efficient and fast mechanism(s) for their long-range dissemination. An interesting possibility concerning this seemingly rapid dispersal comes from the observation that the recombinants were mainly recovered from plums (Table 1). A potential link can be drawn to the success of shanka-tolerant plum cultivars developed in the former Yugoslavia (Rankovic et al., 1994). Since the early 1980s, this new generation of plum cultivars has been deployed, in addition to traditional susceptible cultivars, in Slovenia and other European countries. In that the earliest report of a recombinant isolate can be traced to the former Yugoslavia (Cervera et al., 1993), it is possible that these recombinants have unwittingly been spread through movement of tolerant plum varieties. In addition, the field dispersal of the recombinants may also have been favoured by the absence of eradication efforts. Because PPV accumulates in these cultivars at levels comparable to susceptible cultivars (M. Glasa, unpublished results), aphid-vector spread is efficient.

The data reported here on the biological properties of the recombinant isolates demonstrate that the four isolates analysed, although showing some variability in transmission efficiency, were all able to be aphid-transmitted to a range of susceptible Prunus species. However, with the exception of the Horomerice isolate, the recombinants were only weakly transmitted to peach. These observations agree with previous results (Glasa et al., 2002b) and speculations that PPV recombinants are less adapted than PPV-M isolates to transmission to peach hosts. Another interesting observation is that, despite its T-to-I mutation in the PTK motif, supposed to contribute to binding of the virus capsid protein N-terminal DAG motif to HC-Pro (Blanc et al., 1997), the 143A isolate retained some level of aphid transmissibility. This observation is consistent with the work of Peng et al. (1998) which showed that mutations of this central threonine in the HC-Pro PTK of Zucchini yellow mosaic virus could result in either complete abolition of aphid transmissibility (PAK mutant) or reduced transmission efficiency (PSK and PVK mutants).

In the course of this work, the analysis of all full-length PPV genomic sequences available to date led to the discovery of a previously undetected recombination event with a breakpoint in the 5’ part of the P3 gene, around position 2813. The evidence for the existence of such a recombination comes from several lines of evidence, including the results of the PHYLPRO and GENECONV recombination analyses and the very different contents of the two adjacent regions in phylogenetic signal. The alternative hypothesis, suggesting the existence of different selection pressures on the two regions considered, is not supported by the intra-subgroup diversity values, which are essentially identical between the two regions for both PPV-M and D.

In the 3939 nt long 5’ region, the mean PPV-D/PPV-M divergence level of 3.7% is about twice the corresponding intrasubgroup values (1.6%), which indicates that the recombination event considered is an ancient event, and that following it PPV-D and -M isolates accumulated discriminating mutations. Because one of the parents involved in this event is missing, and because the diversity of PPV-D and PPV-M in the region considered is similar, making it impossible to estimate which is the youngest, it is not currently possible to determine whether the PPV-D or -M subgroup is derived from this ancient event. At present, the only conclusion that can be drawn is that both subgroups share ancestrally the same 5’ NCR-P1-HC-Pro-(Nter)P3 region. This evolutionary history also explains why, whereas the P1 protein has been shown to be one of the most variable regions in potyvirus genomes (Urcuqui-Inchima et al., 2001; Ward et al., 1995), PPV P1, together with the HC-Pro and NCRs, is the most conserved region between the PPV-D and -M subgroups (Palkovics et al., 1993; Fig. 2). On the other hand, when taking into consideration the PPV cherry isolates, which do not share this ancient recombination event, the typical potyvirus variability pattern is restored as P1 is the less conserved gene between the PPV-C and -D or -M isolates (Fig. 2).

It is tempting to speculate about the biological properties conferred on the recombinant isolates by the two recombination events described here. In the case of the PPV-Rec isolates stemming from recombination in the Nlb gene, only limited circumstantial information is available, but these isolates may have retained the limited ability of most PPV-D isolates to infect peach efficiently under field conditions. In the case of the ancient recombination in the P3 gene involving the PPV-D and -M subgroups (and consequently PPV-Rec), the implications are even less clear, but are potentially more important as a number of potyvirus pathogenicity determinants have been identified in...
the recombinated region, including the 5’ NCR (Simon-Buela et al., 1997), P1 (Klein et al., 1994), HC-Pro (Saenz et al., 2001) and P3 (Klein et al., 1994). In particular, in contrast to more classical PPV isolates belonging to the PPV-D and -M subgroups, the PPV-C isolates do not appear to propagate very efficiently in non-cherry Prunus hosts, and it is tempting to hypothesize that this ability is endowed on ‘typical’ PPV isolates by this shared 5’ region. Further experiments aimed at analysing the biological properties of in vitro constructed recombinants between the PPV-C and PPV-D, PPV-Rec or PPV-M isolates should shed light on this question.

Overall, two distinct, broadly successful recombination events have been identified in this study. As a consequence, between one-half and two-thirds of currently known PPV isolates should now be regarded as recombinants, including two of the three major known groups of isolates (PPV-Rec and either PPV-D or PPV-M). These results drastically change our original views of recombinant isolate o6 as an oddity not representative of PPV populations (Cervera et al., 1993) and indicate that, similarly to many other potyviruses (Revers et al., 1996; Glais et al., 2002; Bousalem et al., 2003; Tomimura et al., 2003), recombination has played a major role in the evolution of PPV and in the shaping of its present populations. However, in contrast to the situation observed for Potato virus Y, for example, for which over 10 recombination events have now been identified (Revers et al., 1996; Glais et al., 2002), successful recombination events appear to have been rare in the evolution of PPV. It is possible that this difference is only superficial, but it may also reflect underlying biological differences: although mixed infections of PPV-M and -D have been reported (Candresse et al., 1998; Myrta et al., 1998), recent data concerning the distribution of PPV at the cellular level in mixed infections showed that differently labelled PPV populations appeared to be non-synergistic and remained predominantly separate in the infected plants (Dietrich & Maiss, 2003), which could potentially limit the opportunities for recombination.

ACKNOWLEDGEMENTS

Grateful thanks are expressed to Drs A. Myrta (IAMB Bari, Italy), M. Navratil (Palacky University, Olomouc, Czech Republic), R. Gaborjani (Georgikon Faculty of Veszprem University, Keszthely, Hungary), Ch. Varveri (Benaki Phytopathological Institute, Kifissia, Greece) and A. Stoev (Institute of Plant Protection, Kostinbrod, Bulgaria) for supplying PPV isolates. The work was supported by grant APVT-51-011802 from the Slovak Science and Technology Assistance Agency (APVT) and in part by the VEGA grant 2/3017/23 (Slovak Grant Agency for Science), NAZV QC 1301 (Ministry of Agriculture, Czech Republic) and by OTKA T043388 (Hungary). L. P. was supported by the Janos Bolyai fellowship of the Hungarian Academy of Sciences. We acknowledge the technical assistance of M. Boeglin and J. Novak Nadudvari and thank Kathryn Mayo for helping us with the English of the manuscript.

REFERENCES


