Arch Virol (2005) 150: 1607–1619
DOI 10.1007/s00705-005-0507-z

Archives of Virology
Printed in Austria

Genomic variability of prune dwarf virus as affected by agricultural practice

F. Fonseca, J. D. Neto, V. Martins, and **G. Nolasco**

CDCTPV, Universidade do Algarve, Campus de Gambelas, Faro, Portugal

Received June 29, 2004; accepted January 25, 2005
Published online March 24, 2005

Summary. Twelve new sequences of the coat protein gene of *Prune dwarf virus* (PDV) variants, obtained from almond trees, are presented. Comparison with previously reported sequences of the same region, obtained from other hosts (plum, cherry and peach) revealed not only the existence of a wider range of variants of PDV than formerly predicted, but also the frequent presence of a mixture of variants in each sample. In spite of the heterogeneity found in almond, the amino acid composition of the domain at the N terminus of the coat protein maintained the potential to form an amphipathic helix, and hence the capacity to serve the previously suggested function of binding the viral RNA during particle formation.

Except for synonymous substitutions, measures of nucleotide diversity calculated for the two groups, respectively 13 sequences from almond and 14 sequences from other hosts, were found to be significantly different, with the almond group showing a much higher variability. Analysis of the dendrogram constructed based in all 27 PDV CP sequences did not reveal host specificity, in agreement with previous findings. However, a clear divergence between almond and other hosts sequences could be found. It is discussed that the observed differences between almond and other hosts variants may derive from differences in agricultural practices.

Introduction

Prune dwarf virus (PDV) is a positive-sense RNA plant virus with a tripartite genome that belongs to the genus *Illarvirus* [16], composed of at least 16 distinct viruses, many of which, like PDV, are pathogens of woody plant species. *Illarvirus* spp. have the same genomic organization, encoding functionally similar translation products, as members of the genera *Bromovirus*, *Cucumovirus*, and *Alfamovirus*, all of which belong to the family *Bromoviridae*. RNAs 1 and 2 of viruses in the family *Bromoviridae* are monocistronic and encode the replicase

proteins P1 and P2, respectively. RNA3 is bicistronic and encodes a popypeptide required for cell-to-cell movement (MP) or P3, at the 5′ proximal end, and the viral coat protein (CP) at the 5′ distal end. MP is translated directly from RNA3, while the CP is expressed from a fourth, subgenomic RNA, termed RNA4. This is collinear with the 3′ end of RNA3 and is encapsidated [3].

PDV is one of the most economically important viruses of the genus *Prunus*, frequently occurring in mix infections with the related viruses *Prunus necrotic ringspot virus* (PNRSV) and *Apple mosaic virus* (ApMV). Common biological characteristics, like being transmitted by graft, pollen and seed [6], contribute to their widespread distribution. Symptoms of PDV vary greatly, depending on climate, isolate and host, and range from infections with absence of symptoms to considerable fruit yield reduction. Efforts in developing accurate PDV molecular detection methods have been concerning different authors [20, 8]. Recent molecular characterization [19] of the capsid protein of PDV isolates obtained from different, intensively cultivated hosts (cherry, peach and plum) from Central Europe, made a considerable contribution to the characterization of PDV variants. In that study no reasonable correlation was found between amino acid substitutions and host species and/or geographical origin.

In the case of almond, (*Prunus dulcis*, Miller) information on PDV variants and variability is scarce. In the present state of the art contribution to a better knowledge of PDV can be made by complementing data on CP sequences obtained from cherry, peach and plum with data from almond. In the Algarve region, almond is a traditional crop that has become acclimated in the wild. At present it is exploited as a natural resource, without implementation of irrigation or any agricultural selection criteria. A significant proportion of the plants originate from seed and infection by PDV is prevalent over other Ilarvirus [10]. With the aim of characterizing PDV variants in almond we analyzed isolates obtained from non-selected almond trees in the Algarve region, together with two isolates obtained from the Centre and North of Portugal, and compared them with previously studied isolates from other *Prunus* species.

Material and methods

Infected plant material

One hundred and twenty-eight samples were collected from almond (*Prunus dulcis* Mill.) trees grown in Portugal (south, centre and north). Sampling was done during the more intense growing period (January–March). Young leaves or flower buds were used for testing the presence of PDV by DAS-ELISA, using commercially available polyclonal antybodies, anti PDV IgG and anti PDV IgG conjugated w/AP (Bioreba), according to the manufacturer's instructions.

Isolation of CP gene variants

Positive samples were used for immunocapture followed by reverse transcriptase-polymerase chain reaction (IC/RT-PCR). Approximately 0.2 g of plant material was ground in extraction buffer (100 mM Tris-HCl pH 8.2, 140 mM NaCl, 2% PVP-40, 1% PEG 6000, 0.05% Tween 20) and centrifuged at 5000 × *g* for 10 min. Fifty µl of the supernatant were added to PCR tubes previously coated with anti PDV IgG and incubated overnight (o.n.) at 4 °C. Coating of

PCR tubes was made by adding 50 µl anti PDV IgG 1000× diluted in coating buffer (Na_2CO_3, 15 mM, $NaHCO_3$, 35 mM, pH 9.6) to the tubes. After an o.n. incubation at 4 °C the tubes were washed two times with washing buffer (PBS-Tween, pH 7.4) and one last time with H_2O milliQ sterile. Washing solutions were left in the tubes for 3 min before removal.

After immunocapture, tubes were washed as described above and used for a RT-PCR. The reaction was performed in a total volume of 50 µl containing 10 mM Tris-HCl pH 8.8, 50 mM KCl, 0.08% Nonidet, 4 mM $MgCl_2$, 0.2 mM each dNTP, 0.2 µM each primer, 1 unit of Taq DNA polymerase (Fermentas), 4.65 U of RNase inhibitor (Amersham) and 7.5 U of MuLV reverse transcriptase (Applied Biosystems). RT-PCR was carried out in a Biometra UNO-Thermoblock, and cycling parameters were: a first step of 45 min at 38 °C for first strand cDNA synthesis followed by a cycle of 2 min at 94 °C, 40 cycles of 30 s at 94 °C, 30 s at 50 °C and 50 s at 72 °C, and a final extension of 2 min at 72 °C.

Primers 1175F (5′-CTTGAAGGACCAACCGAG-3′) and 1876R (5′-ATCTGCTAACGC AGGTAAG-3′), previously reported [14], were used for RT-PCR, as well as another pair, combining the upstream primer 1131 (5′-CGATTGGTTAACTCACTTTG-3′) [19], with the downstream primer 1876R. The expected amplified products were 722 bp and 745 bp long respectively, both spanning the last nucleotides of the intergenic region and the whole coat protein open reading frame (ORF2 of RNA3), according to the sequence L28145 available at GenBank.

Amplified products were cloned into pGEM®-T Easy Vector (Promega) as described by the manufacturer and the recombinant plasmid used to transform competent *E. coli* (strain InvαF′–Invitrogen) cells. Recombinant colonies were screened for the insert by PCR with the respective primer pairs.

Single Stranded Conformation Polymorphism (SSCP) analysis

SSCP analysis was performed at two stages of sample characterization: on the amplicons obtained by RT-PCR and on the PCR products obtained from recombinant colonies. Amplicons were denatured in denaturing buffer (formamide: 0.5 M EDTA pH 8: Bromo phenol blue, 95:4:1) at 92 °C for 5 min, followed by 5 min on ice. Samples were loaded on an 8% polyacrylamide (40%, 29:1) gel and electrophoresis was conducted at 4 °C in 1× TBE, at 300 V for 3 hours. Visualization of the patterns obtained was done after silver nitrate staining. The same analysis was conducted after cloning of the RT-PCR amplicons. For this, PCR products of the recombinant colonies, amplified with the respective primer pairs, were used.

Sequencing and sequence analysis

Clones with different SSCP patterns were selected and respective plasmid DNA extracted with Wizard® *Plus* Minipreps DNA Purification System (Promega) according to the manufacturer's instructions. The extracted plasmids were checked for the intactness of the insert by a restriction analysis with *Eco*RI (Fermentas) and commercially sequenced by Macrogen (South Korea). Sequence data were treated using the BioEdit Sequence Alignment Editor program [5]. Multiple alignments were obtained using ClustalX [18]. Phylogenetic statistical analysis was done with MEGA2.1 [7] and statistical analysis was performed in Microsoft®Excel 2002 with StatistiXL add in.

Results

Detection of PDV by DAS-ELISA

The results of the PDV detection by DAS-ELISA, on the 128 samples collected, were taken as an indication of the samples to be further analyzed by IC-RT/PCR.

In this way, 36 samples were considered positive, resulting in a 28% infection rate by PDV in non-selected almond trees.

IC-RT/PCR and SSCP analysis of PDV isolates

Some of the samples could only be amplified with one of the primer pairs used and, in the whole, 17 (47%) out of the 36 samples chosen, based on ELISA results, could not be amplified or the amplified products could not be used for cloning (Fig. 1).

SSCP analysis of the RT-PCR amplicons revealed complex patterns, suggesting the presence of a mixture of variants in each sample. However, subsequent SSCP analysis of recombinant colonies produced simple patterns, generally composed of two bands (Fig. 2), which is in agreement with the expected separation of variants by cloning. Twelve clones with different SSCP patterns, representative of samples from south, centre and north of Portugal, were chosen for sequencing (Table 1). All sequences obtained matched the sequence of the virus coat protein of PDV located between the primers in previously obtained sequences available at GenBank.

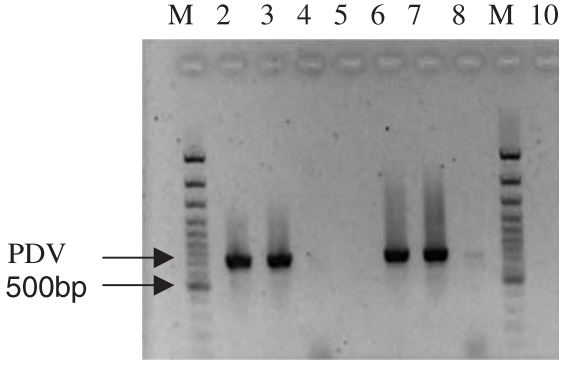

Fig. 1. Gel electrophoresis analysis of IC/RT-PCR products, obtained from field samples. *2*, *3*, *6*, *7* – positive amplification; *4* and *5* – no amplifications; *8* – positive amplification not used for cloning; *10*: negative control of the RT/PCR; *M* – Molecular DNA marker with fragments sizes (bp) of 3000, 2000, 1500, 1200, 1031, 900 to 100 (*M*; 100 bp, Fermentas)

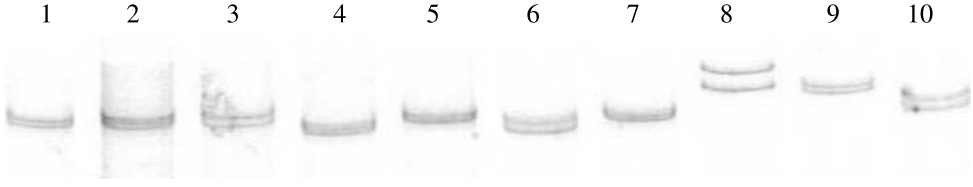

Fig. 2. Example of SSCP patterns obtained from clones of different samples: *1–3*, sample A; *4–7*, sample B; *8–9*, sample C; *10*, sample D

Table 1. Listing and geographical location of PDV variants characterized in this study and available at GenBank

Variant	Host	Location of the tree	Variety/origin of the propagation material (where known)	Accession no.
3.20.1	Almond	Portugal, S.Brás-Estói (South)$^{+}$	–	AY646849
3.20.2	Almond	Portugal, S.Brás-Estói (South)$^{+}$	–	AY646848
3.17.A1	Almond	Portugal, S.Brás-Estói (South)$^{+}$	–	AY646847
3.12N.14	Almond	Portugal, S.Brás-Estói (South)$^{+}$	–	AY646846
3.19.A1.2	Almond	Portugal, S.Brás-Estói (South)$^{+}$	–	AY646845
1.8.A3	Almond	Portugal, Moncarapacho (South)$^{+}$	–	AY646844
4B	Almond	Portugal, Monsanto (Centre)	–	AY646843
26B	Almond	Portugal, Bordeira (South)$^{+}$	–	AY646842
B.11.15	Almond	Portugal, Budens (South)	–	AY646841
cp4.3	Almond	Portugal, Patacão (South)$^{+}$	–	AY646840
cp1.2	Almond	Portugal, Patacão (South)$^{+}$	–	AY646839
N1.2	Almond	Portugal, Foz do Tua (North)	–	AY646838
–	Almond	Portugal, Bordeira (South)$^{+}$	–	AF202117
2/16	Cherry	Czech Republic (North)	Kasinova/Germany	AF208737**
K1	Cherry	Czech Republic (South)	–/–	AF208738**
PS 7/9a	Cherry	Czech Republic (North)	Nanni/–	AF208739**
21/1	Cherry	Czech Republic (North)	Hudson/USA	AF208740**
vanC	Cherry	Czech Republic (North)	Stella compact/Canada	AF208741**
Valtická	Peach	Czech Republic (South–East)	–/–	AF208742**
Nem	Peach	–	–/Germany	AF208743**
Branisovska	Plum	Czech Republic (South)	–/–	AF208744**
Sss	Plum	Czech Republic (South)	–/–	AF208745**
6/54	Plum	Czech Republic (North)	Anna Spath/Germany	AF208746**
7/12	Plum	Czech Republic (North)	Lesidrjanskaja/Russia	AF208747**
–	–	Argentina	–/–	AF317089**
ch 137	?	–	–/–	L28145*
Skierniewice	Cherry	–	–/Poland	U31310

*Bachman et al. [3]; **Vašková et al. [18]; $^{+}$samples collected within a 20 km radius

Table 2. Measures of nucleotide diversity for PDV coat protein gene sequences obtained from almond and from other hosts

Parameter	"Almond"	"Other hosts"	All sequences
Number of sequences	13	14	27
[a]Mean distance within group (±s.e.)	0.070 ± 0.006*	0.050 ± 0.005*	0.069 ± 0.006
[b]Mean distance within group (dS ± s.e.)	0.132 ± 0.013	0.135 ± 0.014	0.155 ± 0.014
[b]Mean distance within group (dN ± s.e.)	0.045 ± 0.006**	0.020 ± 0.004**	0.038 ± 0.005
dN/dS (±s.e.)	0.382 ± 0.022***	0.155 ± 0.007***	0.230 ± 0.013

*$P = 0.01$ (z-test), **$P < 0.001$ (z-test), ***$P = 0.0001$ (Mann-Whitney U-test)
[a]Kimura 2 parameter; [b]Nei-Gojobori, p-distance

```
          1                                                                                                  80
L28145    MSGKAIKSGK PTTRSQSFAL ARKNNNTTPP AGFVKKQ--F PSGSSKSISE WMLHGPNVPV KSFSGMISRT ENLTVNSTAS
U31310    .....A.... ..A...N... ...I..V... ........--. .G.....V.. .......... .......... ..........
3.20.2    ....TT.... .......... ...I.S.... D..IN.KNQI .G....F... .......... .K.P.T.... .....I....
3.17.A1   .F..TT.... .......... ...I...... D..IN.KNQI .G....F... .......... .K.P.T.... .....I....
B.11.15   ...-.T.... ..A....... .......... ........--. .G........ .......... .......... ..........
4B        ..E..T.... ..A....... ...S...... V.......--. SGE..... .......... .R........ ..........
N1.2      .....T.... ..A....... .......A.. ........--. .G........ .......... .N.P..... ..........
3.20.1    ....TT.... ......G... ...I.S.... D..IN.KTQI .G....F... .......... .K.P.T.... ..........
cp4.3     ....TT.... ..A....... ...I..I... DR...N.--I SG.....V.. .......... .T....V... ..........
AF202117  ....TT.... ..A....... ...A.S.... D.......--. .G........ .......... .N....V... ..........
cp1.2     ....I..... ..A....... ...I...... D.S.....--. .G.N..I... .......... .N.P..... ..........
3.19.A1.2 ....PT.... ..A....... ...I...A.. D.L..NKNL. .G.....V.. .......... .N.P..V... ..........
1.8.A3    ....TT.... ..A....... ....K.I... D.S..T.--. .G.....V.. .......... .I.P..V... .....S....
26B       ....PT.... ..A....... ......II.. V.--.N.--P SG.....V.. .T........ .......... ..........
3.12N.14  ....TT.... .......... ...I...... D..---KT-- .G....FV.. .......... .N.P..... ..........
AF208743  .....T.... .......... .......... .......--. .G........ .......... .......... ..........
AF208738  .....T.... .......... .......... V.......--. .G.....T.. .......... .......... ..........
AF208739  .......... .......... .......... .......--. .G........ .......... .......... ..........
AF317089  .-........ .......... .......A.. V.......--. .G.....V.. .......... .N..S..... ..........
AF208737  .......... .......... .......... .......--. .G........ ........S. .......... ..........
AF208744  .......... ..A....... .......... .......--. .G........ .......... ...P...... ..........
AF208742  .......... .......... .......... .......--. .G........ .......... .N........ ..........
AF208740  .....T.... .......... .......... .......--. .G........ ........S. .......... ..........
AF208747  .....N.P.. .......... .......... .......--. .G........ .......... ......V... ..........
AF208746  .....N.P.. .......... .......... .......--. .G........ .......... .R....V... ..........
AF208745  .....T.... .......... .......... .......--. .GE...F... .......... .K........ ..........
AF208741  .....N.... .......... .......... ......R--. .G........ .......... ......V... ..........

          81                                                                                                 160
L28145    GVYYTMKVRE LFKDFAVDTK VYGIVFRYCL DVSNGVYGLI KGFDVNAPVA PNPLQRRKFT AKQASGVQIL APTGMTVGDI
U31310    .......L.D ....L..... .......... .......... .....S.... .......... .......... ........EV
3.20.2    .........D F.....A... .......... .......... .....S.... .......... .......... .....I....
3.17.A1   .A........ .....TA... .......... .......... .....S.... .......... .......... .....I....
B.11.15   ....A..... ......A... .......... .......... .....S.... .......... .......... ..........
4B        ......R... F......... .......... .......... ...N...... .......... .......... ..........
N1.2      .A........ ......T... .......... .......... .....T.... .......... .......... ..........
3.20.1    .......... ......A... .......... .......... .....S.... .......... .......... .....I....
cp4.3     ......R..D F.....A... .......... .......... ....MS.... .......... .......... ..........
AF202117  .......... F...Y..... .......... ...S...... ...N.S.... .......... .......... ..........
cp1.2     .......... .......... .......... .A........ ...N.S.... .......... .......... ..........
3.19.A1.2 .........D F.....A... .......... .......... ...N.S.... .......... .......... ..........
1.8.A3    .......... ......A... .......... .......... ...N...... .........M .......... ..........
26B       .......... .....TA... .......... .......... ...N.S.... ......-... .......... ........S.
3.12N.14  .......... ......A... .......... ...S...... .....S.... .......... .......... ..........
AF208743  .......... ......A... .......... .......... .......... .......... .......... ..........
AF208738  .......... .......... .......... ...S...... .......... .......... .......... ..........
AF208739  .......... ......A... .......... .......... .......... .......... .......... ..........
AF317089  .......... I....PA... .......... .A........ .....S.... .......... .......... ..........
AF208737  .......... .......... .......... .......... .......... .......... .......... ..........
AF208744  .......... .......... .......... .......... .......... .......... .......... ....I.....
AF208742  .......... ......A... .......... ...S...... .......... .......... .......... ..........
AF208740  .......... .....TA... .......... .......... .......... .......... .......... ..........
AF208747  .......... H.....A... .......... .......... .......... .......... .......... ..........
AF208746  .......... H.....A... .......... .......... .......... .......... .......... ..........
AF208745  .......... H.....A... .......... .......... .......... .......... .......... ..........
AF208741  .......... ......A... .......... .......... .....H.... .......... .......... ..........
```

Fig. 3 (*continued*)

```
        161                                                                 222
L28145    PDDLWFVIKY DNAFQPNVPV WFCTQYLQHS MPKRVEVPDS VLYAERDT-A LMDAMDKIVS G*
U31310    .......... .ST...G... .......... ......I... .........-. FI....R... ..
3.20.2    ......G... .....SG... ..G.....Q. ...K..I... .........T. ......R... ..
3.17.A1   .......... .....SG... .......... ......I... .........T. ......R... ..
B.11.15   .......... ..T...D... .......... ......I... .........TL I.....R... ..
4B        .......... ......D... .......... ......I... .........-T .......... ..
N1.2      .......... ..T..SD... .......... ......I... .........T. ......R... ..
3.20.1    .......... .....SG... .......... .....G.... .........P. ......R... ..
cp4.3     .......... ......D... .......... ......I.I. .........T. .......... ..
AF202117  .......V.. ..T...D... .......... ......I.S. .........T. ......R... ..
cp1.2     .......... ......D... .......... ......IS.. .........T. ......R... ..
3.19.A1.2 .......... ......D... .......... ......I.N. .........T. ......R... ..
1.8.A3    .......... ......D... .......... ......I.N. .........T. ......R... ..
26B       .......... ..T...D... .......... ......I... .........TT .L....R... ..
3.12N.14  .......... .....SG... .......... ......I.N. .........T. ......R... ..
AF208743  .......... .......... .......... .......... .........-. .......... ..
AF208738  .......... ......D... .......... .......... .........-. .......... ..
AF208739  .......... ......D... .......... .......... .........-. .......... ..
AF317089  .......... ......DI.. .......... ......I... .........-. .......... ..
AF208737  .......... ......D... .......... .......... .........-. .......... ..
AF208744  .......... ......D... .......... ......I... .........-. .......... ..
AF208742  .......... .......... .......... .......... .........-. .......... ..
AF208740  .......... .......... .......... ......I... .........-. .......... ..
AF208747  .......... ......D... .......... .......... .........-. .......... ..
AF208746  .......... ......D... .......... .......... .........-. .......... ..
AF208745  .......... ......D... .......... .......... .........-. .......... ..
AF208741  .......... ..V...DA.. .......... .......... .........-. .......... ..
```

Fig. 3. Multiple alignment of the deduced amino acid sequences of the coat protein of PDV

Sequences and comparative analysis of the PDV coat protein

Nucleotide sequences, obtained from the isolates mentioned above, were cut to correspond to the coat protein gene, deposited at GenBank (Table 1) and analyzed together with the other PDV coat protein gene sequences available at GenBank. Two groups of sequences were compared: one originating from almond (13 sequences) and one originating from other hosts (peach, cherry and plum; 14 sequences), all available at GenBank. Analysis of the 27 sequences alignment resulted in 666 sites, with 215 (32.3%) variables sites corresponding to 247 nucleotide differences and 116 sites (17.4%) which were parsimony informative. Base frequencies for the "almond" group were estimated as $\pi_{T(U)} = 0.282$, $\pi_C = 0.226$, $\pi_A = 0.265$, $\pi_G = 0.228$, and for other hosts as $\pi_{T(U)} = 0.280$, $\pi_C = 0.227$, $\pi_A = 0.258$, $\pi_G = 0.235$, not showing a statistically significant deviation from equal base composition. Measures of nucleotide diversity were calculated based on all nucleotides (Pi), synonymous substitutions (dS) and non-synonymous substitutions (dN), as shown in Table 2. Except for synonymous substitutions, measures of nucleotide diversity calculated for the two groups were found to be significantly different, with the almond group showing a much higher variability. Both sets showed significant evidence of purifying selection ($P < 0.0001$ under a Mann-Whitney U-test).

Analysis of the amino acid (aa) sequences alignment (Fig. 3) showed an overall mean distance (p-distance, pairwise deletion) of 0.079 (s.e. = 0.010) considering all sites, with 74 aa positions varying within the 222 aa long protein, of which 44

parsimony informative sites. These variations were distributed along the protein sequence, with 45 substitutions in the amino (N) terminal half and 29 in carboxy (C) terminal half. The variations found for the almond group isolates were 65 (of which 38 parsimony informative sites), respectively 40 in the N terminal half and 25 in the C terminal half, whereas for the "other hosts" group, 38 substitutions were found (14 parsimony informative sites), 24 and 14 respectively for the N and C terminal halves. Mean distances within groups where found to be 0.095

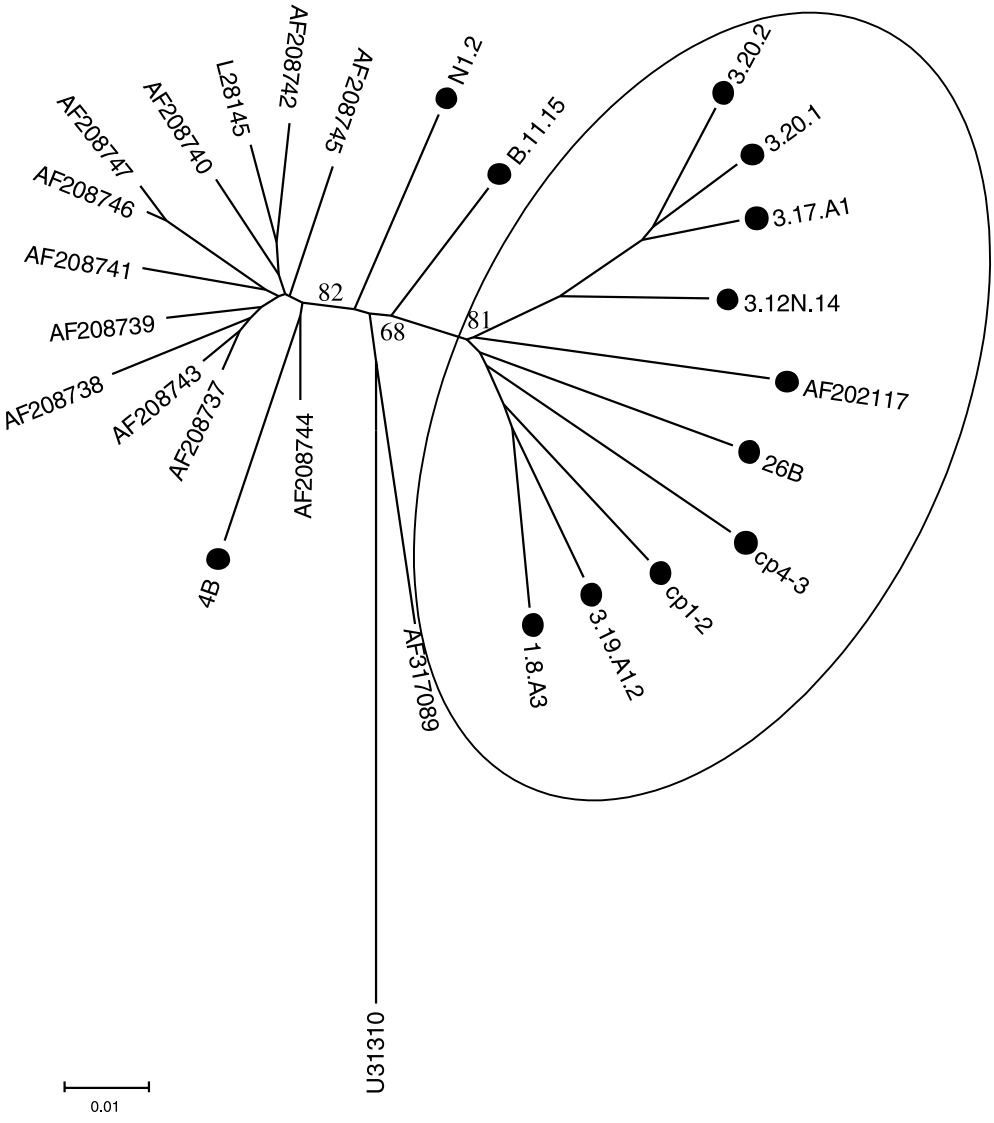

Fig. 4. Dendrogram of PDV coat protein gene sequences characterized in this study (•) and available at GenBank (NJ, 1000 bootstrap, p-distance). Numbers near nodes indicate bootstrap values; sequences obtained from samples collected within a 20 km radius are circled

(s.e. = 0.012) for isolates in the almond host group and 0.043 (s.e. = 0.007) for the "other hosts" group. Mean distance between groups was 0.090 (s.e. = 0.012).

The dendrogram obtained based on nucleotide sequences (Fig. 4), groups together all almond variants originating from samples collected within a 20 km radius (including isolate AF202117 [13]), whereas the remaining three variants are interspersed with isolates from other hosts.

Discussion

This work confirms previous reports of a high prevalence of PDV in almonds, as detected by ELISA [10]. However, a high proportion of positive samples could not be amplified by IC/RT-PCR, even when using two pair of primers. This could be due either to variability in the annealing region of the upstream primers or to unspecific reactions of the antibodies used in ELISA.

We used SSCP to detect different variants prior to sequencing and also for a preliminary assessment of the amplicons variability. The RT-PCR SSCP patterns obtained suggested the isolates were composed by a mixture of variants, which was further confirmed by SSCP analysis of clones and by sequencing. As far as we know, this situation has not been reported for PDV. Interestingly, sequences previously presented by other authors [19] were obtained by direct sequencing of amplified products, which would have been difficult to perform if a mixture of variants had also been present in their samples. Apparently, either mixtures were not detected by those authors, or are characteristic of the kind of samples used in the present work.

The PDV RNA3 is 2129 nucleotides long and contains two large open reading frames (ORFs), separated by an intergenic region of 72 nucleotides [3]. The first ORF was found to encode for the movement protein (882 nucleotides), while the second ORF was suggested to encode for the capsid protein (657 nucleotides). Deducing from the protein length in aminoacids reported previously [19], sequences of at least 654 nucleotides were considered to encompass the length of the PDV CP gene. In our case, variable lengths were obtained after sequences were cut at the beginning of ORF2, with the longest reaching 666 nt. Observation of the alignment obtained in this study indicated two regions, respectively between 80–120 nt and 620–630 nt, as the most variables between isolates from almond and from other hosts. Isolates from almond showed not only a higher score of amino acids substitutions, but also some isolates analyzed in this study had two more aminoacids in positions 38 and 39 and one more in position 209. The amino acid composition of the domain at the N terminus of the PDV coat protein when compared to related virus [3] like *Alfalfa mosaic virus* (AIMV), is equally enriched in basic amino acids and, when the region is folded into a helical conformation, three basic amino acids are nearly all located on one side, forming an amphipathic helix. This suggests that the same function of binding the viral RNA during particle formation may be served by PDV coat protein [3]. In our case, and in spite of the heterogeneity found in almond PDV isolates, the same potential to form an

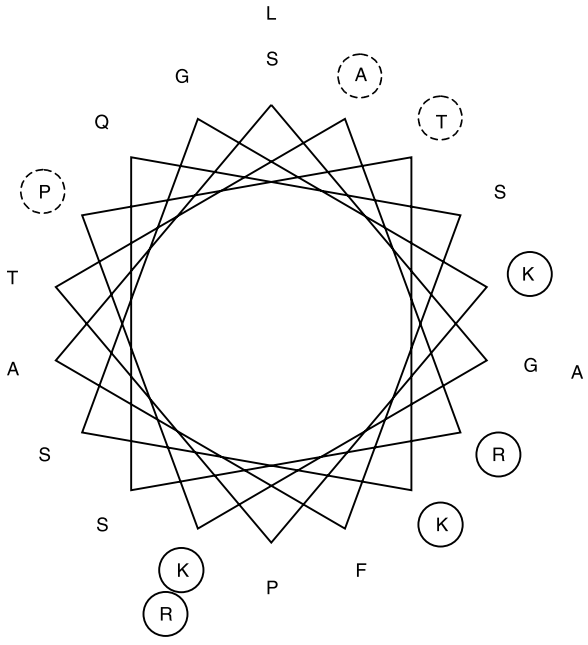

Fig. 5. Helical wheel representation of amino acids 2–22 in variant 3.19.A1.2 coat protein. The wheel shows the projection of all the side chain positions along the helix axis onto a plane. Like in the helical wheel representation presented previously [3] the basic amino acids are primarily localized to one side of the helix (hydrophilic face) and are circled; amino acids circled with a dashed line differ from the ones found by those authors

amphipathic helix was observed for all isolates, from which an example is shown in Fig. 5.

Studies on related ilar- and alfamoviruses coat proteins indicate that one of the arginines in the N-terminal part of the CP is crucial for RNA-binding activity [17, 21]. Previously it had been shown [1] that the amino acid residues surrounding this crucial arginine were well conserved in the majority of viruses of this group. In PDV sequences from other hosts [19] the arginine is located at position 14 of amino acids 10–19 in a KPTT**R**SQSFA context and variability of PDV observed in this region is limited. In our work, the crucial arginine was also located at position 14, in a KPTT**R**SQSFA or KPTA**R**SQSFA context. Only one of the "almond group" isolates, 3.20.1, had a G_{17}, instead of the prevailing S_{17}. Amino acids P_{11}, S_{15} and A_{19}, previously associated with the RNA-binding activity of the CP [1], were conserved in all PDV isolates considered in this study.

In a similar study with PNRSV isolates [2], no clear correlation with host specificity could be found. Working with PDV, Vašková et al. [19] also did not find any reasonable correlation between amino acid substitutions and host species and/or geographical origin. As such we considered all sequences previously characterized by those authors as one single group. In this way, we were able to compare a similar number of sequences: 13 obtained from almond and 14 from other hosts. Analysis of nucleotide diversity as presented in Table 2 showed similar value for synonymous substitutions for both sets. On the contrary, when non-synonymous substitutions were considered, a significantly higher value ($P < 0.0001$, Mann-Whitney U-test) was found for almond variants. This data seems to indicate a significant difference in the type or extent of selection each group is subjected to, with implications to their diversification. Apparently a more intense

purifying selection pressure is operating on the "other hosts" group. Chare and Holmes [4] analyzed selection pressures in the capsid genes of plant RNA viruses and found that vector borne viruses are subjected to a greater purifying selection than non-vectored viruses. Power [12] also suggested that surface interactions involved in vector mediated transmission impose a selective constraint to the capsid protein. Similarly, one would expect constraints against CP variability in other transmission mechanisms that depend on surface interactions. However, in the case of PDV, the study of cellular location of the virus by *in situ* RT-PCR [15] showed the viral particles inside pollen cells or ovules. In this case pollen or seed transmission operating would not impose more selective constraints to the viral surface than the general virus-plant host interaction or graft transmission. Thus the differences in purifying selection pressures between the almond and "other hosts" group should not be attributable to the mechanism of transmission. As the "other hosts" variants come from cultivated orchards we hypothesize that the differences in selection pressure are related to the agricultural practice. Eventually some less aggressive variants could have been unintentionally selected by farmers and maintained through graft transmission. Interestingly, the work of Ooi and Yahara [11] on the genetic variation of geminivirus in sexual and asexual host plant populations has suggested that amino acid replacements of ORF C4, a gene known to function as a host range determinant, are accelerated in the sexual population.

The dendrogram constructed (Fig. 4), showed that some PDV almond variants are interspersed with PDV variants from other hosts. However there is a clear divergence between the PDV variants obtained in Algarve and the remaining. All the variants that were collected from an area which has an approximate radius of 20 km cluster in the same branch of the dendrogram while maintaining the level of non-synonymous nucleotide diversity (0.041, s.e. = 0.005) found for almond variants. This is probably due to close spatial relationships and facilitated genetic flux via pollen exchange, as graft transmission of PDV is excluded due to the kind of agricultural practice.

In conclusion, we also did not find host species-related specificity in the variants obtained in our study. However, it becomes obvious that, when the 13 sequences from almond are included, PDV CP diversity substantially increases, especially at the aa level. From the data here analyzed it appears that, in cultivated *Prunus*, certain variants of the PDV are selected in relation to the phenotypic selection the hosts are being subjected to, leading to less variability. Also, when transmission is done by grafting, through vegetative propagation material, geographic speciation is probably precluded. Based on the observed differences between the two groups, we suggest that modern agricultural practice tends to decrease PDV genomic diversity.

Acknowledgements

This work was supported by the grant POCTI/AGR-39095/2001, from Fundação para a Ciência e a Tecnologia (FCT).

Note added in proof

During February 2005 a set of six new PDV coat protein gene sequences from Hungarian isolates were made available at GenBank by Krizbai and Nemeth, respectively AY554278, AY554277 and AY554275 from sweet cherry, AY554274 and AY551441 from sour cherry and AY554276 from *Prunus avium*. The dendrogram obtained after addition of those new sequences is in concurrence with the one shown in Fig. 4. reinforcing the suggestion put forward in the discussion section.

References

1. Ansel-McKinney P, Scott SW, Swanson M, Ge X, Gehrke L (1996) A plant viral coat protein RNA binding consensus sequence contains a crucial arginine. EMBO J 15: 5077–5084
2. Aparicio F, Myrta A, Di Terlizzi B, Pallás V (1999) Molecular variability among isolates of prunus necrotic ringspot virus from different *Prunus* spp. Phytopathology 89: 991–999
3. Bachman EJ, Scott SW, Xin G, Vance VB (1994) The complete nucleotide sequence of Prune Dwarf Ilarvirus RNA3: Implications for coat protein activation of genome replication in Ilarviruses. Virology 201: 127–131
4. Chare ER, Holmes EC (2004) Selection pressure in the capsid genes of plant RNA viruses reflect mode of transmission. J Gen Virol 85: 3149–3157
5. Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41: 95–98
6. Kelley RD, Cameron HR (1986) Location of prune dwarf and prunus necrotic ringspot viruses associated with sweet cherry pollen and seed. Phytopathology 76: 317–322
7. Kumar S, Tamura K, Jacobsen IB, Nei M (2001) MEGA, Molecular Evolutionary Genetics Analysis, version 2.1. http://www.megasoftware.net
8. Mekuria G, Ramesh SA, Alberts E, Bertozzi T, Wirthensohn M, Collins G, Sedgley M (2003) Comparison of ELISA and RT-PCR for the detection of Prunus necrotic ring spot virus ans prune dwarf virus in almond (*Prunus dulcis*). J Virol Methods 114: 65–69
9. Nei M, Kumar S (2000) Molecular evolution and phylogenetics, 1st edn. Oxford University Press, New York
10. Nolasco G, Neves MA, Faria EA (1991) Distribuição no Algarve de virus do grupo Ilarvirus em amendoeira e suas consequências na produção. 1ª Aproximação. Rev Cienc Agrar 14: 33–37
11. Ooi K, Yahara T (1999) Genetic variation of geminiviruses: comparison between sexual and asexual host plant populations. Mol Ecol 8: 89–97
12. Power AG (2000) Insect transmission of plant viruses: a constraint on virus variability. Curr Opin Plant Biol 3: 336–340
13. Raquel MH, Miguel CM, Nolasco G, Oliveira MM (2000) Construction and testing of a transformation vector to introduce PDV resistance in almond. Acta Horticult 521: 73–82
14. Raquel H, Tereso S, Oliveira MM, Nolasco G (1999) Establishing a protocol for IC-RT-PCR detection of prune dwarf virus in almond. Petria 9: 153–156
15. Silva C, Tereso S, Nolasco G, Oliveira MM (2003) Cellular location of prune dwarf virus in almond sections by In Situ reverse transcription-polymerase chain reaction. Phytopathology 93: 278–285
16. Rybicki EP (1995) Family Bromoviridae. In: Murphy FA, Fauquet CM, Bishop DHL, Ghabrial SA, Jarvis AW, Martelli GP, Mayo MA, Summers MD (eds), Virus

taxonomy. Sixth Rep. Int. Comm. Taxon. Viruses. Springer-Verlag, Wien New York, pp 450–457
17. Swanson MM, Ansel-McKinney P, Houser-Scott F, Yusibov V, Loesch-Fries LS, Gehrke L (1998) Viral coat protein peptides with limited sequence homology bind similar domains of alfalfa mosaic virus and tobacco streak virus RNAs. J Virol 72: 3227–3234
18. Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The ClustalX windows interface: flexible strat egies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 24: 4876–4882
19. Vašková D, Petrzik K, Špak J (2000) Molecular variability of the capsid protein of the prune dwarf virus. Eur J Plant Pathol 106: 573–580
20. Youssef SA, Shalaby AA, Mazyad HM, Hadidi A (2002) Detection and identification of prune dwarf virus and plum pox virus by standard and multiplex RT-PCR probe capture hybridization (RT-PCR-ELISA). J Plant Pathol 84: 113–119
21. Yusibov V, Loesch-Fries LS (1998) Functional significance of three basic N-terminal amino acids of alfalfa mosaic virus coat protein. Virology 208: 405–407

Author's address: Filomena Fonseca, Universidade do Algarve, CDCTPV, Campus de Gambelas, 8005-139 Faro, Portugal; e-mail: ffonseca@ualg.pt