



Molecular analysis of prune dwarf virus reveals divergence within non-Turkish and Turkish viral populations

Adyatma Irawan Santosa¹ · Ali Çelik² · Miroslav Glasa³ · Çiğdem Ulubaş Serçe⁴ · Filiz Ertunç⁵

Received: 7 April 2021 / Accepted: 9 May 2023 / Published online: 30 May 2023
© The Author(s) under exclusive licence to Società Italiana di Patologia Vegetale (S.I.Pa.V.) 2023

Abstract

Prune dwarf virus (PDV) infection affects the production of stone fruits (*Prunus* spp.) around the world, including in Turkey. In this study, we analyzed the available nucleotide (nt) and amino acid (aa) sequences of P1, P2, movement protein (MP), and coat protein (CP) genes to assess the molecular variability and structure of non-Turkish and Turkish PDV populations. Global isolates were clustered into three major groups in nt-based phylogenetic trees of MP and CP. Comparison of geographically different isolates showed that the MP aa sequences were more conserved than the CP aa sequences. The major aa changes in MP and CP showed the most effect on group 3 isolates. MP and CP-based analyses revealed the lowest number of variable sites (S), total number of mutations (η), average number of nt differences between sequences (k), and nt diversity (per site) (π) values in group 3 isolates, indicating high genomic similarities. The Turkish isolates manifested the highest S and η , and the second highest k and π values were observed in the comparison with CP, confirmed the high intra-group divergence of the Turkish isolates belonging to the three phylogroups. A stronger purifying selection pressure was observed in the MP region than in the CP region. Neutrality tests deduced negative values for all the MP and CP phylogroups and positive values for the P1, P2, Bulgarian, and Polish populations. MP and CP comparisons showed the divergence of the three phylogroups, as revealed by the results of the K_S^* , K_{ST}^* , Z^* , and S_{nn} permutation statistical tests. The fixation index (F_{ST}) values suggested that non-Turkish PDV isolates had a higher genetic distance compared with Turkish isolates from group 3 and the Turkish isolates from the other two groups.

Keywords Identity percentages · *Illarvirus* · Genetic linkage · Molecular variation · Population structure · Selection pressure

✉ Adyatma Irawan Santosa
adyatma.i.s@ugm.ac.id
Ali Çelik
alichelik@ibu.edu.tr
Miroslav Glasa
Miroslav.Glasa@savba.sk
Çiğdem Ulubaş Serçe
cigdemserce@ohu.edu.tr
Filiz Ertunç
ertunc@agri.ankara.edu.tr

¹ Department of Plant Protection, Faculty of Agriculture, Universitas Gadjah Mada, Jl. Flora No. 1, Sleman 55281, Yogyakarta, Indonesia
² Department of Plant Protection, Faculty of Agriculture, Bolu Abant İzzet Baysal University, Bolu 14030, Turkey
³ Biomedical Research Center of the Slovak Academy of Sciences, Institute of Virology, Dúbravská Cesta 9, Bratislava 84505, Slovakia
⁴ Department of Plant Production and Technologies, Faculty of Agricultural Sciences and Technologies, Niğde Ömer Halisdemir University, Niğde 51240, Turkey
⁵ Department of Plant Protection, Faculty of Agriculture, Ankara University, Diskapi, Ankara 06110, Turkey

Introduction

Prune dwarf virus (PDV, genus *Ilarvirus*, family *Bromoviridae*) is a commonly found pathogen in stone fruits, such as sour cherry (*Prunus cerasus*), sweet cherry (*P. avium*), almond (*P. dulcis*), apricot (*P. armeniaca*), plum (*P. domestica*), and peach (*P. persica*) (Vaskova et al. 2000; Paduch-Cichał et al. 2011). PDV plant infection is recognized through various symptoms, including leaf deformity, chlorotic rings and spots on leaves, and stunting. These symptoms can vary depending on the host plants, virus isolates, and environmental factors (Fonseca et al. 2005; Predajna et al. 2017; Kamenova et al. 2019). Except for the transmission by grafting, the transmission of PDV occurs vertically via pollen or seeds (Kelley and Cameron 1986; Silva et al. 2003). As a result, PDV infection highly affects the health and yields of susceptible plants (Uyemoto and Scott 1992).

A multipartite PDV genome consists of three positive-sense single-stranded RNA segments. RNA1 possesses a single open reading frame (ORF) that encodes the non-structural P1 protein, which is involved in virus replication (Pallas et al. 2013). Similarly, RNA2 contains a single ORF encoding the P2 protein, which is an RNA-dependent RNA polymerase and an element of the replication complex (Kozieł et al. 2017). RNA3 has two ORFs: ORF3a and ORF3b. ORF3b encodes the movement protein (MP), which is immediately translated from RNA3. In ORF3a, a downstream subgenomic promoter region enables the transcription of sgRNA4, which encodes ORF3b from which the viral coat protein (CP) is translated (Pallas et al. 2013). The MP generates tubular structures that allow cell-to-cell viral transportation via plasmodesmata (Kozieł et al. 2018); on the other hand, CP performs an important role in genome activation, which is essential for viral replication and encapsidation of viral particles (Bachman et al. 1994; Pallas et al. 2013).

Initial investigations on variations in the PDV genome in different *Prunus* spp. mostly focused on the CP gene (Vaskova et al. 2000; Öztürk and Çevik 2015; Zong et al. 2015; Kamenova et al. 2019; Çelik and Ertunç 2020). These studies indicated the high divergence of the nucleotide (nt) sequences of the CP gene from different hosts (Fonseca et al. 2005; Ulubaş Serçe et al. 2009; Öztürk and Çevik 2015). In spite of the variable biological and serological characteristics of PDV isolates, the molecular characterization of their CP sequences failed to clearly differentiate isolates on the basis of host or geographical origin relationship (Vaskova et al. 2000; Öztürk and Çevik 2015). Thus, molecular variability studies must be extended to other coding regions, e.g., the MP gene (Predajna et al. 2017).

Selection analysis can be used to reveal evolutionary processes to gain insights into the molecular epidemiology

and extent of virus genomic variability (Farzadfar and Pourrahim 2019; Santosa and Ertunç 2021). The genetic diversity analysis of plant virus populations can also be used to determine the best disease management options, including engineering plants that show resistance (Sokhandan-Bashir and Melcher 2012). Although PDV, which is distributed worldwide, is a common and traditional stone fruit virus, its population genetic structure and molecular evolution remain poorly understood and require further exploration. This study aimed to determine the genetic variability of PDV populations using the most recent database data and compare the population genetic structures of isolates of different geographical origins, with emphasis on the PDV population in Turkey.

Materials and methods

PDV isolates and recombination analysis

The nt and amino acid (aa) sequences of P1, P2, MP, and CP genes of PDV, which are available in the National Center for Biotechnology Information (NCBI) GenBank, including 41 Turkish isolates originating from four *Prunus* spp. from different regions of the country (Ulubaş Serçe et al. 2009; Öztürk and Çevik 2015; Çelik and Ertunç 2020), were analyzed in this study. The complete nt sequence of the P1 gene spans nt 39–3206 in RNA1 (reference isolate: accession No. NC_008039). The P2 gene spans nt 32–2398 in RNA2 (reference isolate: accession No. NC_008037). The MP spans nt 260–1141 in RNA3, and the CP gene at nt 1214–1870 in RNA3 (reference isolate: accession No. NC_008038). Some divergent isolates show differences in the length of their respective coding regions due to indel mutations.

A total of 51 isolates with complete RNA3 sequences, which are available in the NCBI GenBank database (accessed on October 11, 2021), were aligned and trimmed using CLUSTAL W implemented in MEGA X v.10.2.4 (Kumar et al. 2018) to create two alignments involving the MP and CP regions. A total of 26 other isolates, for which only complete or nearly complete MP sequences were available, were added to the MP region alignment to generate a final MP dataset of 77 isolates. Then, 103 other isolates, for which only complete CP sequences were available, were added to the CP region alignment to generate a final CP dataset of 154 isolates. Two other alignments were performed on 9 and 11 isolates for which the complete ORF of P1 and P2, respectively, are available in NCBI GenBank (Online Resource 1-Table 1).

The recombination signals on the P1, P2, MP, and CP sequence alignments were scanned using Recombination Detection Program (RDP v.4.56) (Martin et al. 2015);

credible signals included only those supported by five or more of the RDP, GENECONV, BootScan, MaxChi, ChiMaera, SiScan, and 3Seq algorithms with a Bonferroni-corrected P value of <0.05 . Further analyses of this study excluded all recombinant isolates to conduct an accurate phylogenetic analysis.

Phylogenetic and comparative nt and aa analyses

Four phylogenetic trees based on MP and CP nt sequences were built using maximum-likelihood (ML) and neighbor-joining (NJ) algorithms in MEGA X v.10.2.4 (Kumar et al. 2018). Two ML trees were constructed to observe the phylogrouping of isolates based on the P1 and P2 coding regions. The models for nt substitution of each alignment were selected using the lowest Bayesian information criterion score. The statistical significance of isolate clusters was determined using Tamura 3-parameter (Tamura 1992), with 1000 bootstrap replicates, uniform rate among sites, and complete deletion of missing data treatment.

Analysis of aa variations in the MP and CP regions

Estimation of the percentage identities of P1, P2, MP, and CP genomic regions at the nt and aa levels was performed using Sequence Demarcation Tool v1.2 software (Muhire et al. 2014). The aa sequences of the MP region of 8, 7, and 5 isolates represented groups 1, 2, and 3, respectively. The aa sequences of the CP region of 19, 11, and 10 isolates represented groups, 1, 2, and 3. All these aa sequences were aligned in BioEdit v.7.2.5 (Hall 1999) to examine the sequence variation among isolates from different phylogroups.

Genetic diversity and polymorphism analyses

The genetic diversities of P1, P2, MP, and CP nt sequences among non-Turkish and Turkish isolates were estimated using DnaSP v.6.12.03 (Rozas et al. 2017), based on the confidence intervals of several parameters: the number of haplotypes = h , haplotype diversity = H_d , the number of variable sites = S , total number of mutations = η , average nt differences between sequences = k , nt diversity (per site) = π ; the ratio of nonsynonymous to synonymous site (dN/dS) = ω .

Neutral selection analysis

Statistical tests of Tajima's D , Fu and Li's D^* , and Fu and Li's F^* (without out-group; window length: 100 sites and step size: 25 sites) were performed using DnaSP v.6.12.03 (Rozas et al. 2017) to explain the demography of PDV

populations. The differences between the number of segregating sites and the average nt differences served as the basis for Tajima's D (Tajima 1989). The differences between the number of singletons and the total number of mutations were the basis for Fu and Li's D^* , and that for Fu and Li's F^* included the differences between the number of singletons and the average nt differences between pairs of sequences (Fu and Li 1993).

Gene flow and genetic differentiation among populations

The K_S^* , K_{ST}^* , Z^* , S_{nn} , and F_{ST} values (Hudson et al. 1992; Hudson 2000), which determined the genetic differentiation of P1, P2, MP, and CP nt sequences among phylogroups, respectively, were calculated using DnaSP v.6.12.03 (Rozas et al. 2017). The K_{ST}^* value is close to zero in the absence of genetic divergence among the tested populations (Tsompana et al. 2005). The small Z^* value indicates a low genetic isolation (Hudson et al. 1992). The value of S_{nn} ranges from 0.5 (the exact same population) to 1 (distinctly separated populations) (Hudson 2000). The F_{ST} value ranges between a minimum of 0, which is used to describe strictly the same populations, to a maximum of 1, which is used to describe entire distinct populations (Hudson et al. 1992; Tsompana et al. 2005). $F_{ST} > 0.25$ usually implies a high gene flow and large genetic isolation among populations (Gao et al. 2016). The calculated F_{ST} values among the three sampling regions in Turkey (AHK region: Adana, Hatay, and Kahramanmaraş; BB region: Bursa and Bilecik; AI region: Afyon and Isparta) provided insights into the genetic distances of isolates from different regions in Turkey.

Results

Phylogenetic and recombination analyses

The exact same topology was observed between the two NJ and two ML phylogenetic trees based on MP and CP alignments, that is, the PDV isolates were divided into three major lineages (Fig. 1; Online Resource 2- Figs. 1 and 2). Turkish isolates were detected in all the three phylogroups. DL, NS9, and PCH4 isolates were consistently positioned in the same groups in the trees based on the P1, P2, MP, and CP comparisons. However, given the currently limited availability of the complete sequences of P1 and P2, a general consensus for the phylogrouping of the four genomic regions could not be established (Fig. 1; Online Resource 2-Table 1). Recombinant events were detected in three different parts of P1 of KY883327 (Table 1).

Illustrations

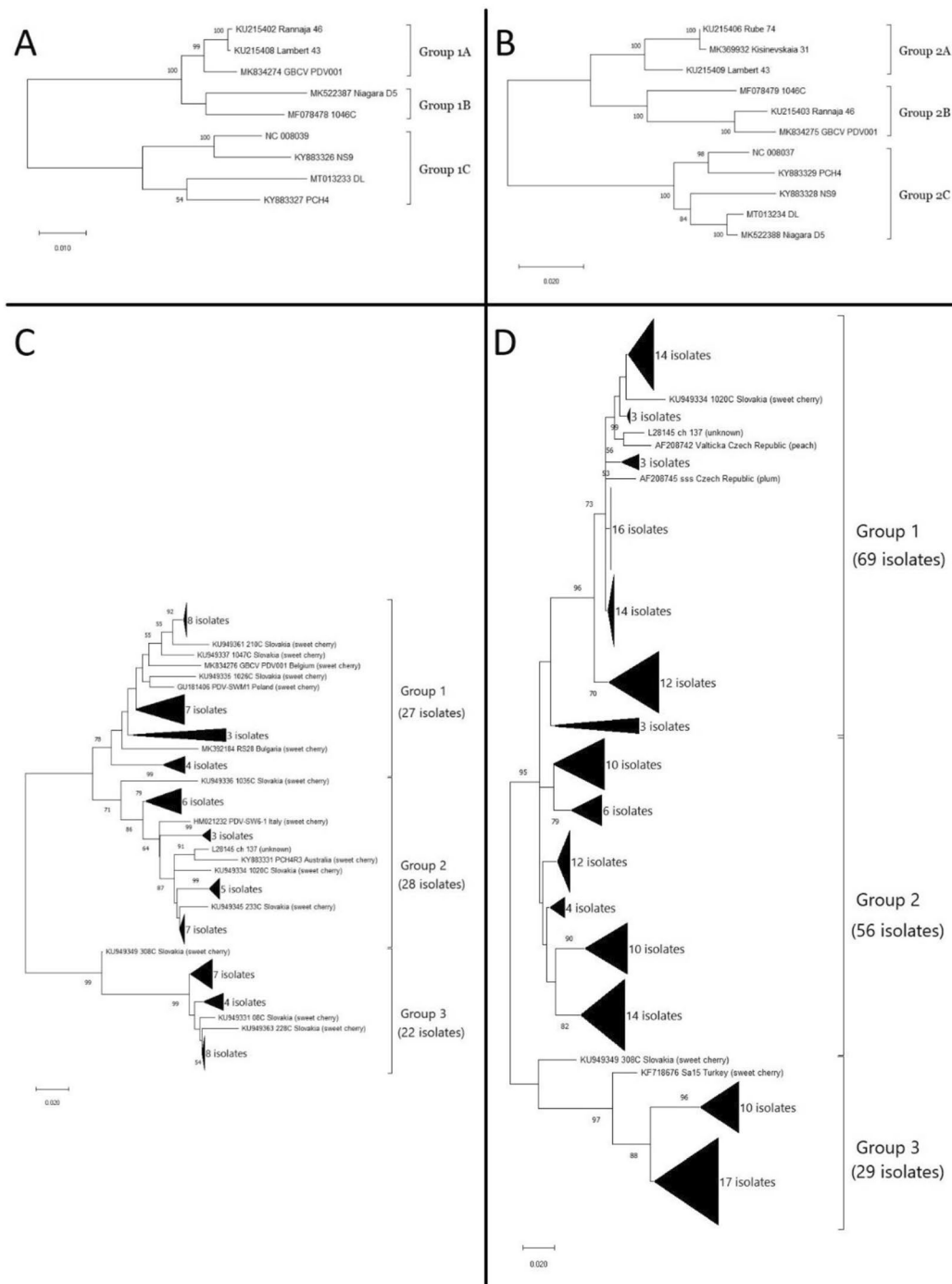


Fig. 1 Maximum-likelihood trees constructed using Tamura 3-parameter model (T92) with uniform rates and complete deletion of missing data implemented in MEGA X software and based on nucleotide sequences of prune dwarf virus P1 (A), P2 (B), MP (C), and CP (D) genomic regions. A total of 77 and 154 isolates were included in the

MP and CP comparison, respectively. Bootstrap values on each branch were supported with 1000 bootstrap replications (only values > 50% are shown). The accession numbers for isolates in the compressed branches are listed in Online Resource 1-Table 1

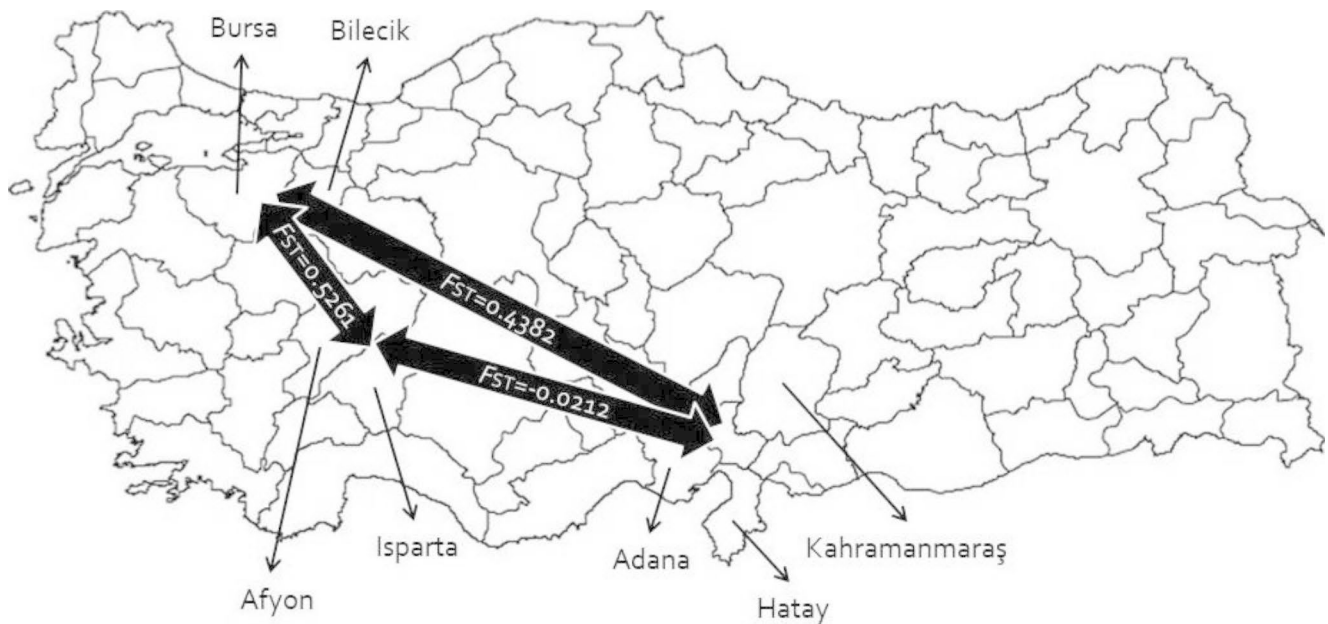


Fig. 2 The most significant genetic linkages between the prune dwarf virus populations of three sample collection sites in Turkey: BB region: Bursa+Bilecik provinces (11 isolates); AI region: Afyon+Isparta provinces (25 isolates); AHK region: Adana+Hatay+Kahramanmaraş

provinces (4 isolates). Linkages are indicated by their fixation index (F_{ST}) values which were calculated based on coat protein region comparisons

Table 1 Putative recombination events detected in the P1 region by RDP4 analysis of PDV isolates

No.	Recombinant	Parents: major/minor	Breakpoints ¹ (start/end)	RDP-implemented method ² (<i>P</i> value)
1.	KY883327 (Australia)	MT013233 (China)/NC_008039 (Canada)	540–1050	R (5.732×10^{-07}) G (8.442×10^{-06}) B (1.159×10^{-01}) M (3.373×10^{-06}) C (1.698×10^{-04}) S (1.468×10^{-10}) 3 S (1.505×10^{-03})
2.	KY883327 (Australia)	MT013233 (China)/NC_008039 (Canada)	1430–1702	R (6.798×10^{-07}) G (1.531×10^{-05}) B (1.000×10^{-03}) M (1.936×10^{-05}) C (1.534×10^{-05}) S (1.202×10^{-07})
2.	KY883327 (Australia)	MT013233 (China)/NC_008039 (Canada)	2698–3012	R (1.298×10^{-04}) G (1.451×10^{-05}) B (1.753×10^{-01}) C (3.160×10^{-02}) S (2.639×10^{-06}) 3 S (3.105×10^{-03})

¹ Position in alignment

² R=RDP; G=GENECOV; B=BootScan; M=MaxChi; C=Chimaera; S=Siscan; 3 S=3Seq

Analysis of nt and aa variations

The comparative analysis suggested the relatively high nt and aa identities among all studied isolates in the P1, P2, MP, and CP regions. In general, the nt identity among isolates in the MP region was higher than the aa identity, and

in the CP region, the identity was higher at the aa level (Table 2).

The MP and CP genomic regions exhibited highly conserved aa sequences. The three phylogroups in the MP region revealed five major aa residue mutations at positions 14, 189, 258, 279, and 280 (RefSeq NC_008038) (Online Resource 2-Fig. 3). The CP region showed seven

Table 2 Nucleotide (nt) and amino acid (aa) identities in the P1, P2, movement (MP) and coat protein (CP) genomic regions among PDV phylogroups

Phylogroup	Identity (%)		Phylogroup	Identity (%)		Phylogroup	Identity (%)			
	P1 (n=9)			P2 (n=11)			MP (n=77)		CP (n=154)	
	nt	aa		nt	aa		nt	aa	nt	aa
All isolates	91.4–99.8	95.4–99.9	All isolates	87.5–99.8	86.8–99.9	All isolates	86.2–100	89.7–100	84–100	79.8–100
1 A (n=3)	98.7–99.8	99.4–99.9	2 A (n=3)	97.1–99.8	97.8–99.9	1 (n=27 and 69, respectively)	92.7–100	94.6–100	88.9–100	82.6–100
1B (n=2)	96.3	98.7	2B (n=3)	93.7–97.7	95.6–98.1	2 (n=28 and 56, respectively)	91.6–100	91.2–100	87.8–100	86.7–100
1 C (n=4)	94.3–97.4	96.8–98.5	2 C (n=5)	94.3–99.2	95.3–98.9	3 (n=22 and 29, respectively)	93.8–99.6	95.4–100	90–100	89–100

major aa residue changes at positions 6, 13, 24, 27, 97, 173, and 177 (Online Resource 2-Fig. 4). In addition, six additional residues were observed in the nt sequences of five almond isolates from the Portuguese population (3.19.A1.2, 3.12.N.14, 3.17.A1, 3.20.2, and 3.20.1), which translated into two additional aa residues at positions 38–39. Twelve Portuguese, one Italian almond, and one apricot isolate from Greece (B3) showed three additional nts (ACT) at positions 622–624, which were translated into one extra aa residue (T, Threonine) at position 208. These changes were not observed in the other two almond isolates (4B from Portugal and PDV-AL2 from Italy) (Online Resource 2-Fig. 4). More substitutions were observed in the aa sequences of the MP and CP genes of group 3 isolates compared with the other group isolates.

Genetic diversity and polymorphism analyses

In general, the three phylogroups and different countries showed higher nt diversities (π values) among in the CP genomic region than in the MP genomic region, except the populations from Bulgaria, Poland, and Slovakia, which exhibited higher diversities in the MP region. Only the Hd value of group 3 in the MP genetic comparison reached 1.000, which further supported the indication of a relatively low nt diversity among PDV isolates. The Turkish population had Hd and π values of 0.998 and 0.0806, respectively. In the CP genomic region, the Turkish population had considerably higher S , η , and k values than the populations from other countries. These results showed that Turkish isolates, which were distributed in all three phylogroups, shared a large genetic variation. MP and CP gene comparisons showed that group 2 had higher S , η , and k values than the other two phylogroups (Table 3).

The dN/dS was also calculated to analyze the evolutionary forces acting on the MP and CP genes. According to the results, the MP and CP genomic regions experienced very intense negative selection pressures, with a stronger pressure observed in the MP region ($\omega = 0.0669$) than in the CP

region ($\omega = 0.1795$). The purifying selection pressures on the MP region of the Bulgarian, Slovak, and Polish isolates ($\omega = 0.0701$, 0.0747, 0.0811, respectively) were also more intense than on their counterparts in the CP genomic region ($\omega = 0.1314$, 0.1478, 0.1487). Therefore, the MP gene may have played an important role in the adaptive evolution of global PDV isolates. As expected, the negative selection pressures on the highly variable CP region of Turkish isolates ($\omega = 0.2279$) was weaker than those of other countries except Portugal ($\omega = 0.3366$) (Table 3). In addition, P2 had a higher genetic divergence among its isolates and experienced a weaker purifying pressure than P1 (Table 3).

Neutral selection analysis

Neutrality tests estimated positive values for P1 and P2 and negative values for all three phylogroups in the MP and CP genomic regions. However, the calculations by Fu and Li's D^* and Fu and Li's F^* tests for all groups 1 and 2 isolates were more convincing given the support from significant values (Table 4). All three neutrality tests predicted positive values for Bulgarian and Polish populations in the MP region comparison. As for the CP gene, Fu and Li's D^* and Fu and Li's F^* tests only assigned positive values to the Polish population (Table 4).

Gene flow and genetic differentiation among populations

The analysis showed statistically significant K_S^* , K_{ST}^* , and Z^* values in the MP and CP genetic comparisons of all phylogroups. The S_{nn} values among phylogroups all reached 1.0000 except group 1/group 2 for the CP gene (0.9760). These results were in line with those of the phylogenetic analysis, which showed significant genetic variations among the three PDV lineages. In the comparison of Turkish isolates with other isolates from different phylogroups, non-significant K_S^* , K_{ST}^* , and Z^* values were obtained only in All¹/Turkey 1 and group 1¹/Turkey 1 (Table 5). However,

Table 3 Genetic diversity and polymorphism analyses of P1, P2, MP and CP genomic regions of PDV variants from different phylogroups and countries

Phylogroups	N	h	Hd	S	η	k	π	dS	dN	ω
P1	9	9	1.000	475	159	193.194	0.0609	0.2143	0.0149	0.0695
P2	11	11	1.000	529	576	205.218	0.0868	0.2621	0.0378	0.1442
MP	77	73	0.999	263	325	54.946	0.0769	0.2721	0.0182	0.0669
Group 1	27	25	0.994	162	171	30.202	0.0422	0.1453	0.0119	0.0819
Group 2	28	26	0.995	167	190	34.712	0.0484	0.1601	0.0154	0.0962
Group 3	22	22	1.000	105	113	18.610	0.0259	0.0889	0.0077	0.0866
CP	154	144	0.999	369	528	51.798	0.0788	0.2106	0.0378	0.1795
Group 1	69	64	0.997	256	318	33.095	0.0504	0.1272	0.0267	0.2099
Group 2	56	52	0.997	262	318	39.703	0.0604	0.1526	0.0317	0.2077
Group 3	29	28	0.998	139	156	30.034	0.0457	0.1162	0.0234	0.2014
Countries (MP)										
Bulgaria	14	13	0.989	101	106	33.703	0.0471	0.1654	0.0116	0.0701
Poland	10	9	0.978	164	184	67.933	0.0770	0.2664	0.0216	0.0811
Slovakia	40	40	1.000	220	270	58.165	0.0811	0.2851	0.0213	0.0747
Countries (CP)										
Bulgaria	13	12	0.987	75	81	23.064	0.0351	0.1043	0.0137	0.1314
Czech Rep.	13	11	0.962	117	123	27.923	0.0425	0.1227	0.0177	0.1443
Hungary	5	5	1.000	111	115	53.800	0.0819	0.2214	0.0383	0.1729
Italy	8	8	1.000	108	119	39.464	0.0601	0.1658	0.0272	0.1641
Poland	23	18	0.976	160	184	49.336	0.0751	0.2139	0.0318	0.1487
Portugal	13	13	1.000	153	172	43.282	0.0659	0.1337	0.0450	0.3366
Slovakia	23	23	1.000	168	195	50.692	0.0771	0.2199	0.0325	0.1478
Turkey	41	39	0.998	241	295	52.961	0.0806	0.1957	0.0446	0.2279
Turkey 1	3	3	1.000	39	39	26.000	0.0396	0.1158	0.0161	0.1391
Turkey 2	20	19	0.995	181	201	37.442	0.0569	0.1206	0.0372	0.3085
Turkey 3	18	17	0.993	95	102	20.778	0.0316	0.0697	0.0197	0.2826

N: number of isolates, h: number of haplotypes, Hd: haplotype diversity, S: number of variable sites, η : total number of mutations, k: average number of nucleotide differences between sequences, π : nucleotide diversity (per site), dN: non-synonymous nucleotide diversity, dS: synonymous nucleotide diversity, ω : dN/dS

Turkey 1, 2, 3: Turkish isolates in Group 1, 2, and 3, respectively

these results were driven only by the presence of three Turkish isolates in group 1.

Compared with the CP region, the MP region showed higher F_{ST} values, which were greater than 0.25, among phylogroups. The analysis results also indicated that in general, the gene flow from groups 1 and 2 to group 3 was lower than that from group 1 to group 2. Furthermore, the gene flow from non-Turkish to Turkish isolates in group 3 (Turkey 3) was lower than the gene flow to Turkish isolates in groups 1 (Turkey 1) and 2 (Turkey 2) (Table 5). Pairwise comparisons with two other collection sites revealed that the isolates collected in BB region had F_{ST} values > 0.25 (AI region: Afyon and Isparta provinces; AHK region: Adana, Hatay, and Kahramanmaraş provinces) (Table 5). Figure 2 summarizes graphically the genetic linkages between isolates from the three sample collection sites in Turkey.

Discussion

The significant genetic diversity of RNA plant viruses is largely due to recombination and mutations in virus populations. The distribution of variability itself depends on two main evolutionary processes: genetic drift and selection (García-Arenal et al. 2001). Knowledge of the molecular basis for adaptations of plant RNA viruses to different hosts and geographical ranges and their population structure can provide data for the development of effective detection tools and strategies for efficient phytosanitary measures (Jones 2009; Çelik and Ertunç 2021). PDV is one of the important viruses hampering stone fruit production worldwide, including that in Turkey. Therefore, this study performed updated and detailed diversity and evolutionary analyses of P1, P2, MP, and CP genes of non-Turkish and Turkish isolates currently available in GenBank.

Our study showed the absence of recombination events in the MP and CP genomic regions, which is in agreement with the lack of recombination in the CP region

Table 4 Results from demographic test statistics between sequences of P1, P2, MP and CP genomic regions of PDV populations

Phylogroups	Fu and Li's D^*	Fu and Li's F^*	Tajima's D
P1	0.06887 ns	0.07532 ns	0.06107 ns
P2	0.25192 ns	0.27437 ns	0.21125 ns
MP	-1.66226 ns	-1.45697 ns	-0.58589 ns
Group 1	-1.77807 ns	-1.89569 ns	-1.25356 ns
Group 2	-1.27024 ns	-1.44600 ns	-1.13007 ns
Group 3	-2.09610 ns	-2.28090 ns	-1.61553 ns
CP	-3.35046 **	-2.92900 *	-1.47102 ns
Group 1	-2.53919 *	-2.66218 *	-1.75146 ns
Group 2	-2.77620 *	-2.74424 *	-1.52826 ns
Group 3	-1.40147 ns	-1.47780 ns	-0.94548 ns
Countries (MP)			
Bulgaria	0.37292 ns	0.32642 ns	0.04975 ns
Poland	0.19082 ns	0.22449 ns	0.22171 ns
Slovakia	-0.31190 ns	-0.83653 ns	-0.77183 ns
Countries (CP)			
Bulgaria	-0.16101 ns	-0.29753 ns	-0.52783 ns
Czech Rep.	-1.33849 ns	-1.53660 ns	-1.34964 ns
Hungary	-0.11024 ns	-0.13791 ns	-0.19293 ns
Italy	-0.65403 ns	-0.75924 ns	-0.76400 ns
Poland	0.75053 ns	0.58858 ns	-0.04195 ns
Portugal	-1.07201 ns	-1.20725 ns	-1.00457 ns
Slovakia	-0.05924 ns	-0.10818 ns	-0.16395 ns
Turkey	-1.83935 ns	-1.77127 ns	-0.86237 ns
Turkey 1	nd	nd	nd
Turkey 2	-2.10299 ns	-2.20962 ns	-1.40829 ns
Turkey 3	-1.48278 ns	-1.64635 ns	-1.25782 ns

ns = not significant

nd = not determined, a minimum of four data sets are needed in the calculation

* $0.01 < P \text{ value} > 0.05$; ** $0.001 < P \text{ value} > 0.01$; *** $P \text{ value} < 0.001$

Turkey 1, 2, 3: Turkish isolates in Groups 1, 2, and 3, respectively

reported by Kalinowska et al. (2014). These results were also in agreement with the observation of Predajňa et al. (2017), who reported that the recombination breakpoints of isolates PDV-PE247 (accession No. GU187047), PDV-PA78 (GU181403), PDV-SOF15P11 (GU181404), 1048 C (KU949338), and 235 C (KU949346) were found on the central part of RNA3 (within or close to the intergenic region) and did not affect the MP and CP genes directly.

Phylogenetic analysis of the studied PDV isolates showed their affiliation to three major groups (1, 2, and 3) and the same topology in MP- and CP-based trees. The phylogroups did not exhibit a distinct relation with hosts and geographical origins, and the isolates obtained from different *Prunus* spp. and countries were integrated together in the same group and subgroup. This finding indicates the long-term diversification of PDV and the spread of PDV isolates facilitated by the intensive global trade of natural PDV hosts. This phylogrouping was similar to those in other recent reports (Öztürk and Çevik 2015; Predajňa et al. 2017; Kamenova et al. 2019; Song et al. 2020), which indicated that genetically, PDV isolates are not highly variable, as shown by the results

of the identity analysis of nt and aa sequences conducted in this study. Most isolates whose MP and CP sequences are available were clustered in an identical manner to the MP and CP trees, as also observed by Predajňa et al. (2017). Hence, these data raised the possibility that phylogrouping of PDV isolates can be based on either the MP or CP gene nt sequences, with the same results. However, further analysis involving P1 and P2 showed that substantially more complete sequences are needed to fully understand the PDV phylogeny.

MP aa sequences were highly conserved among the isolates, suggesting the production of synonymous substitutions by most of nt changes in the region. On the other hand, the CP aa sequences presented considerably lower identity percentages than the nt sequences. The majority of nt changes in the CP region generated nonsynonymous aa substitutions, most of which occurred on group 3 isolates, which contained seven major aa differences compared with groups 1 and 2 isolates (Online Resource, Fig. 2). Thus, nt changes became one of the bases for PDV phylogrouping. However, these changes showed no negative effect

Table 5 Genetic differentiation estimates for lineages of PDV, based on MP and CP gene sequence comparison

Comparison	$^aK_S^*$	$^aK_{ST}^*$	<i>P</i> value	$^aZ^*$	<i>P</i> value	<i>Snn</i>	<i>P</i> value	$^bF_{ST}$
MP								
Group 1(n=27)/ Group 2(n=28)	3.4029	0.0808	0.0000***	5.6456	0.0000***	1.0000	0.0000***	0.3941
Group 1(n=27)/ Group 3(n=22)	3.1241	0.1634	0.0000***	5.3372	0.0000***	1.0000	0.0000***	0.6721
Group 2(n=28)/ Group 3(n=22)	3.2238	0.1574	0.0000***	5.3673	0.0000***	1.0000	0.0000***	0.6665
CP								
Group 1(n=69)/ Group 2(n=56)	3.5058	0.0564	0.0000***	7.4987	0.0000***	0.9760	0.0000***	0.2926
Group 1(n=69)/ Group 3(n=29)	3.3798	0.1024	0.0000***	6.8869	0.0000***	1.0000	0.0000***	0.5531
Group 2(n=56)/ Group 3(n=29)	3.4951	0.1001	0.0000***	6.5205	0.0000***	1.0000	0.0000***	0.5246
Non-Turkey $^{\lambda}$ (n=113)/ Turkey(n=41)	3.7849	0.0229	0.0000***	8.1882	0.0000***	0.9468	0.0000***	0.1284
Group 1 $^{\lambda}$ (n=66)/ Turkey(n=41)	3.5785	0.0519	0.0000***	7.2905	0.0000***	0.9327	0.0000***	0.2539
Group 2 $^{\lambda}$ (n=36)/ Turkey(n=41)	3.6584	0.0457	0.0000***	6.6894	0.0000***	0.9610	0.0000***	0.2105
Group 3 $^{\lambda}$ (n=11)/ Turkey(n=41)	3.6409	0.0545	0.0000***	5.8678	0.0000***	1.0000	0.0000***	0.3839
Non-Turkey(n=113)/ Turkey 1(n=3)	3.7695	-0.0012	0.8090 ns	7.8145	0.7910 ns	0.9638	0.2710 ns	0.0838
Non-Turkey(n=113)/ Turkey 2(n=20)	3.7409	0.0169	0.0000***	7.9415	0.0000***	0.9699	0.0000***	0.1714
Non-Turkey(n=113)/ Turkey 3(n=18)	3.6577	0.0528	0.0000***	7.6656	0.0000***	1.0000	0.0000***	0.4954
Group 1 $^{\lambda}$ (n=66)/ Turkey 1(n=3)	3.4311	-0.0021	0.7430 ns	6.7750	0.7160 ns	0.9391	0.3050 ns	-0.0074
Group 2 $^{\lambda}$ (n=36)/ Turkey 2(n=20)	3.4982	0.0303	0.0000***	6.1549	0.0000***	0.9464	0.0000***	0.1427
Group 3 $^{\lambda}$ (n=11)/ Turkey 3(n=18)	2.8619	0.1249	0.0000***	4.4387	0.0000***	1.0000	0.0000***	0.5004
Turkey 1(n=3)/ Turkey 2(n=20)	3.5179	0.0289	0.0000***	4.3415	0.0000***	1.0000	0.0000***	0.3719
Turkey 1(n=3)/ Turkey 3(n=18)	2.8696	0.1056	0.0000***	4.0942	0.0000***	1.0000	0.0000***	0.6553
Turkey 2(n=20)/ Turkey 3(n=18)	3.2113	0.1496	0.0000***	4.8363	0.0000***	1.0000	0.0000***	0.6025
Turkish isolates origin$^{\tau}$								
AHK(n=4)/BB(n=11)	2.4013	0.1959	0.0000***	3.3032	0.0000***	0.9333	0.0050**	0.4382
AHK(n=4)/AI(n=25)	3.8413	0.0022	0.3030 ns	5.0128	0.4110 ns	0.8276	0.2180 ns	-0.0212
BB(n=11)/AI(n=25)	3.3206	0.1198	0.0000***	4.8954	0.0000***	0.9722	0.0000***	0.5261

ns = not significant

* 0.01 < *P* value > 0.05; ** 0.001 < *P* value > 0.01; *** *P* value < 0.001 $^aK_S^*$, $^aK_{ST}^*$, *Z* and *Snn* are test statistics of genetic differentiation $^bF_{ST}$, coefficient of gene differentiation, which measures inter-population diversity λ = minus Turkish isolates τ Group of provinces where Turkish isolates were originated: AHK: Adana, Hatay, and Kahramanmaraş; BB: Bursa and Bilecik; AI: Afyon and Isparta

Turkey 1, 2, 3: Turkish isolates in group 1, 2, and 3, respectively

on the fitness of group 3 isolates. The aa substitution A₁₃T in KPTTRSQSFA motif in the CP (position aa 10–19) of some isolates, as reported by Ulubaş Serçe et al. (2009) and Predajna et al. (2017), was also observed on a number of

other isolates belonging to all three phylogroups; this finding suggested that T₁₃ and A₁₃ residues are interchangeable in this motif, which is involved in the RNA-binding activity of CP. As mutation evidently contributed to shaping of the

population structure of PDV, the information obtained here can be considered in breeding programs to develop varieties with increased resistance to PDV.

Only the Hd value of group 3 in MP and CP comparisons among phylogroups reached 1.000, further indicating that the PDV population consisted of many identical haplotypes. The isolates from Slovakia (in MP and CP) and Hungary, Italy, and Portugal (in CP) also had Hd values of 1.000, showing the high divergences of their respective isolates. In the MP and CP comparisons, the lowest S , η , k , and π values were all correlated to group 3 isolates, indicating their high genomic similarities. Among the populations from various countries, Turkish isolates had the highest S and η and the second highest k and π values (after Hungarian) in the CP comparison. Thus, the Turkish isolates, which constituted almost a third of the tested isolates and were positioned in all the three phylogroups, exhibited a high variability. The high variations were probably due to the weak purifying selection pressure on Turkish isolates, especially those in group 2 (Turkey 2) ($\omega = 0.3085$). Turkey 2 isolates also recorded higher S , η , k , and π values than Turkey 1 and 3. The stronger negative constraint on the MP than the CP gene is likely the most important evolutionary factor that made the MP gene more conserved than the CP gene.

Neutrality tests assigned positive values to P1, P2, Bulgarian, and Polish populations, and these values indicated that the populations emerged from an excess of intermediate frequency alleles and can evolve out of other population through bottlenecks, balancing, and/or structural selection (Biswas and Akey 2006). The positive values also suggested that Bulgarian and Polish isolates emerged or were introduced a long time ago and have been endemic in their respective country, or that they were results of vegetative propagation from the same source(s). Moreover, members of the Polish population, which consisted of apricot, peach, plum, sour cherry, and sweet cherry isolates, might have infected different hosts at various times. On the other hand, the negative values for the other PDV populations belonging to different phylogroups and countries indicated their population expansion (Tsompana et al. 2005). Therefore, new haplotypes can be predicted to emerge among these expanding populations.

Gene flow and genetic differentiation analyses presented significant K_S^* , K_{ST}^* , Z^* , and S_{nn} values and vividly distinguished the three phylogroups in the MP and CP comparisons. Based on the obtained F_{ST} values, group 1 had a higher gene flow to group 2 than group 3 and thus shared more genetic similarities with group 2 than with group 3. This finding is in agreement with phylogenetic analyses that clustered group 1 closer to group 2 than group 3. The F_{ST} values also indicated that Turkey 1 was more similar to Turkey 2 than Turkey 3. Moreover, gene flow analysis

revealed that non-Turkish isolates had considerably higher differences from Turkey 3 ($F_{ST} = 0.4954$) than from Turkey 1 and 2 ($F_{ST} = 0.0838, 0.1714$, respectively).

The observation of gene flow among isolates from three different collection sites in Turkey showed that the 11 isolates from BB region, which were all from peach and formed a distinct subcluster within group 3, were genetically isolated from the other Turkish isolates obtained in either AI or AHK region, which were from apricot, sweet cherry, and sour cherry and were distributed in all three phylogroups. Meanwhile, the comparison between isolates from AI and AHK regions yielded a very small negative F_{ST} value (-0.0212) and suggested the lack of divergence between them. Therefore, the obtained data also indicated relatively strong relations between the phylogrouping of Turkish isolates with host species and geographic locations.

In summary, the detailed genetic diversity and evolutionary studies on P1, P2, MP, and CP genes have expanded our knowledge on the structure of non-Turkish and Turkey PDV populations. These results may provide a basis for studies of other genomic regions, including isolates from other parts of the world (Asia, Africa, and Oceania) from which very few isolates have been molecularly characterized to date.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42161-023-01412-2>.

Authors' contributions All authors contributed to the study conception and design. AIS prepared sequence alignments, performed molecular and evolutionary analyses, interpreted the results, and wrote the manuscript. AÇ performed phylogenetic and evolutionary analysis and wrote the manuscript. MG interpreted the results and revised the manuscript. ÇUS revised the manuscript. FE supervised the analyses. All authors read and approved the final manuscript.

Funding AIS was supported by the grant (3612/UN1/PN/PN/PT.01.03/2022) from the Faculty of Agriculture, Universitas Gadjah Mada. MG was supported by the grant (VEGA 2/0030/20) from the Scientific Grant Agency of the Ministry of Education and Slovak Academy of Sciences.

Data Availability The authors declare that the data supporting the findings of this study are included in this published article and/or its supplementary materials. The datasets generated and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Ethics approval This article does not contain any studies with human participants nor animals performed by any of the authors.

References

- Bachman EJ, Scott SW, Xin GE, Vance VB (1994) The complete nucleotide sequence of prune dwarf ilarvirus RNA 3: implications for coat protein activation of genome replication in ilarviruses. *Virology* 201(1):127–131. <https://doi.org/10.1006/viro.1994.1272>
- Biswas S, Akey JM (2006) Genomic insights into positive selection. *Trends Genet* 22(8):437–446. <https://doi.org/10.1016/j.tig.2006.06.005>
- Çelik A, Ertunç F (2020) Bursa ve Bilecik illeri şeftali yetiştiriciliği yapılan alanlarda *prune dwarf virus* (PDV) yaygınlığı ve genetik çeşitliliği / distribution and genetic diversity of *prune dwarf virus* (PDV) in peach growing areas in Bursa and Bilecik provinces (in turkish). *Int J Agric Wildl Sci* 6(1):66–74. <https://doi.org/10.24180/ijaws.663272>
- Çelik A, Ertunç F (2021) Reverse transcription loop-mediated isothermal amplification (RT–LAMP) of plum pox potyvirus Turkey (PPV–T) strain. *J Plant Dis Prot* 128(3):663–671. <https://doi.org/10.1007/s41348-021-00447-w>
- Farzadfar S, Pourrahim R (2019) Positive selection and recombination shaped the large genetic differentiation of *Beet black scorch virus* population. *PLoS ONE* 14(4):e0215574. <https://doi.org/10.1371/journal.pone.0215574>
- Fonseca F, Neto JD, Martins V, Nolasco G (2005) Genomic variability of prune dwarf virus as affected by agricultural practice. *Arch Virol* 150(8):1607–1619. <https://doi.org/10.1007/s00705-005-0507-z>
- Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics* 133(3):693–709
- Gao F, Lin W, Shen J, Liao F (2016) Genetic diversity and molecular evolution of arabis mosaic virus based on the CP gene sequence. *Arch Virol* 161(4):1047–1051. <https://doi.org/10.1007/s00705-015-2729-z>
- García-Arenal F, Fraile A, Malpica JM (2001) Variability and genetic structure of plant virus populations. *Annu Rev Phytopathol* 39:157–186. <https://doi.org/10.1146/annurev.phyto.39.1.157>
- Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* 41:95–98
- Hudson RR (2000) A new statistic for detecting genetic differentiation. *Genetics* 155:2011–2014
- Hudson RR, Boos DD, Kaplan NL (1992) A statistical test for detecting geographic subdivision. *Mol Biol Evol* 9:138–151. <https://doi.org/10.1093/oxfordjournals.molbev.a040703>
- Jones RA (2009) Plant virus emergence and evolution: origins, new encounter scenarios, factors driving emergence, effects of changing world conditions, and prospects for control. *Virus Res* 141(2):113–130. <https://doi.org/10.1016/j.virusres.2008.07.028>
- Kalinowska E, Mroczkowska K, Paduch-Cichal E, Chodorska M (2014) Genetic variability among coat protein of *prune dwarf virus* variants from different countries and different *Prunus* species. *Eur J Plant Pathol* 140(4):863–868. <https://doi.org/10.1007/s10658-014-0502-x>
- Kamenova I, Borisova A, Popov A (2019) Incidence and genetic diversity of *prune dwarf virus* in sweet and sour cherry in Bulgaria. *Biotechnol Biotechnol Equip* 33(1):980–987. <https://doi.org/10.1080/13102818.2019.1637278>
- Kelley RD, Cameron HR (1986) Location of prune dwarf and prunus necrotic ringspot viruses associated with sweet cherry pollen and seed. *Phytopathology* 76(3):317–322
- Kozielec E, Bujarski JJ, Otulak K (2017) Molecular biology of *prune dwarf Virus*—A lesser known member of the Bromoviridae but a vital component in the dynamic virus–host cell interaction network. *Int J Mol Sci* 18(12):2733. <https://doi.org/10.3390/ijms18122733>
- Kozielec E, Otulak-Kozielec K, Bujarski JJ (2018) Ultrastructural analysis of *prune dwarf virus* intercellular transport and pathogenesis. *Int J Mol Sci* 19(9):2570. <https://doi.org/10.3390/ijms19092570>
- Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Mol Biol Evol* 35(6):1547–1549. <https://doi.org/10.1093/molbev/msy096>
- Martin DP, Murrell B, Golden M, Khoosal A, Muhire B (2015) RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol* 1:1–5. <https://doi.org/10.1093/ve/vev003>
- Muhire BM, Varsani A, Martin DP (2014) SDT: a virus classification tool based on pairwise sequence alignment and identity calculation. *PLoS ONE* 9(9):e108277. <https://doi.org/10.1371/journal.pone.0108277>
- Öztürk Y, Çevik B (2015) Genetic diversity in the coat protein genes of *prune dwarf virus* isolates from sweet cherry growing in Turkey. *Plant Pathol J* 31(1):41–49. <https://doi.org/10.5423/PPJ.OA.07.2014.0063>
- Paduch-Cichal E, Sala-Rejczak K, Mroczkowska K, Boscia D, Poter O (2011) Serological characterization of prune dwarf virus isolates. *J Plant Prot Res* 51(4):389–392. <https://doi.org/10.2478/v10045-011-0063-3>
- Pallas V, Aparicio F, Herranz MC, Sanchez-Navarro JA, Scott SW (2013) The molecular biology of ilarviruses. *Adv Virus Res* 87:139–181. <https://doi.org/10.1016/b978-0-12-407698-3.00005-3>
- Predajňa L, Sihelská N, Benediková D, Šoltys K, Candresse T, Glasa M (2017) Molecular characterization of *prune dwarf virus* cherry isolates from Slovakia shows their substantial variability and reveals recombination events in PDV RNA3. *Eur J Plant Pathol* 147(4):877–885. <https://doi.org/10.1007/s10658-016-1055-y>
- Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol Biol Evol* 34:3299–3302. <https://doi.org/10.1093/molbev/msx248>
- Santosa AI, Ertunç F (2021) Phylogenetic and diversity analyses of Garlic common latent virus based on the TGB and CP gene sequence. *Plant Prot Sci* 57(3):179–187. <https://doi.org/10.17221/149/2020-PPS>
- 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(3):278–285. <https://doi.org/10.1094/phyto.2003.93.3.278>
- Sokhandan-Bashir N, Melcher U (2012) Population genetic analysis of grapevine fanleaf virus. *Arch Virol* 157(10):1919–1929. <https://doi.org/10.1007/s00705-012-1381-0>
- Song S, Zhang L, Wang Q, Zhang J-hua, Li Z-nan (2020) Identification and characterization of the first complete genome sequence of prune dwarf virus in China. *Phytopathol Mediterr* 59(2):371–376. <https://doi.org/10.14601/Phyto-11492>
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123(3):585–595
- Tamura K (1992) Estimation of the number of nucleotide substitutions when there are strong transition-transversion and G + C-content biases. *Mol Biol Evol* 9(4):678–687. <https://doi.org/10.1093/oxfordjournals.molbev.a040752>
- Tsompana M, Abad J, Purugganan M, Moyer J (2005) The molecular population genetics of the *Tomato spotted wilt virus* (TSWV) genome. *Mol Ecol* 14:53–66. <https://doi.org/10.1111/j.1365-294X.2004.02392.x>
- Ulubaş Serçe Ç, Ertunç F, Öztürk A (2009) Identification and genomic variability of *prune dwarf virus* variants infecting stone fruit trees in Turkey. *J Phytopathol* 157(5):298–305. <https://doi.org/10.1111/j.1439-0434.2008.01486.x>

- Uyemoto JK, Scott SW (1992) Important diseases of Prunus caused by viruses and other graft-transmissible pathogens in California and South Carolina. *Plant Dis* 76(1):5–11
- Vaskova D, Petrzik K, Spak J (2000) Molecular variability of the capsid protein of the prune dwarf virus. *Eur J Plant Pathol* 106(6):573–580. <https://doi.org/10.1023/A:1008742513754>
- Zong X, Wang W, Wei H, Wang J, Yan X, Hammond RW, Liu Q (2015) Incidence of sweet cherry viruses in Shandong province, China and a case study on multiple infection with five viruses. *J Plant Pathol* 97(1):61–68. <https://doi.org/10.4454/JPP.V97I1.047>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.