# Molecular analysis of the complete genomes of apricot and plum isolates of *Plum pox virus* detected in a *Prunus* germplasm collection in Almaty's Pomological Gardens, Kazakhstan

D. James[1,a], D. Sanderson[1], S. Dolgikh[2] and S. Spiegel[3]

[1]Sidney Laboratory – Centre for Plant Health, Canadian Food Inspection Agency, 8801 East Saanich Road, North Saanich, British Columbia, V8L 1H3, Canada; [2]Kazakhstan Research Institute of Horticulture and Viticulture, Almaty, Kazakhstan; [3]Department of Plant Pathology, Agricultural Research Organization, The Volcani Center, Bet Dagan 50 250, Israel.

***Abstract***

**Partial molecular characterization of isolates of *Plum pox virus* (PPV) from Kazakhstan was described by Spiegel et al. (2004). The isolates were classified as members of the Dideron (D) strain based on: a) RFLP analysis of a 243-bp amplicon in the region of the C terminus of the coat protein (CP) fragment, b) RFLP analysis of an 836-bp amplicon in the P3-6K$_1$ coding region and c) analysis of the CP coding region. The plum isolates were detected in plum trees maintained in the plum selection-orchard of the Pomological Garden, Almaty, Kazakhstan. The apricot isolates were detected in wild apricot seedlings obtained originally from Zailiyski Alatou at the northern edge of the Tien-Shan Mountains, but maintained in the Pomological Garden, Almaty, for about 15 years. Disease symptoms were observed on fruit from one of these infected apricot trees and on the leaves of the infected plum trees. PPV was never detected in any wild apricot plants growing in their natural habitat in Zailiyski Alatou. In this recent study, the virus genomes were sequenced completely and their analysis confirmed their identities as members of the D strain of PPV. Interestingly, in some cases, the genomes of the apricot and plum isolates were very similar, with identities of approximately 99.7% (nucleotide and amino acid sequence). Phylogenetic analyses seem to indicate a close relationship of the Kazakhstan apricot and plum isolates to each other, and to the PPV isolate BIII/2 (GenBank accession no. GU461890), a plum isolate from Slovakia. The Kazakhstan isolates and the Slovakia isolate are approximately 99% identical (nucleotide and amino acid sequence). The deletion event described by Spiegel et al. (2004) in the plum isolate P5R8 (KZPLa1) was confirmed.**

**Keywords:** Kazakhstan, apricot germplasm, PPV D, complete genomes, deletion event

## INTRODUCTION

Apricot (*Prunus armeniaca*) trees growing in the Zailiyski Alatou area at the northern edge of the Tien-Shan Mountains are considered a valuable source of germplasm for breeding programs that may be interested in developing cold-tolerant/frost-resistant and/or disease-resistant cultivars (Spiegel et al., 2004). This area belongs to the Central Asian ecogeographical group of *P. armeniaca*, and is considered to be among the oldest and richest in its biodiversity (Kostina, 1969). This represents valuable germplasm that should be protected and preserved. Unfortunately, this valuable resource is threatened by cattle grazing and by the cutting down of trees.

To assist with the preservation of this germplasm, collections of apricot seedlings obtained originally from the Zailiyski Alatou area were established in the Pomological

[a]E-mail: Delano.James@inspection.gc.ca

Garden, Almaty, Kazakhstan. Over many years, the Almaty Pomological Gardens has gathered an extensive collection of *Prunus* germplasm from around the world, including over 200 varieties of plum. Several years ago, virus-like symptoms were observed on the leaves of some plum trees and on fruit from one apricot tree. Spiegel et al. (2004) used a range of molecular analysis and characterization tools, including protein (serology)-based and nucleic acid-based assays, to confirm the presence of *Plum pox virus* (PPV) in plum and apricot samples from the Pomological Gardens. Strain typing that included RFLP analysis targeting a 243-bp fragment in the coat protein (CP) coding region (Wetzel et al., 1991), an 836-bp fragment in the P3-6$K_1$ region (Glasa et al., 2002) and sequencing of these RT-PCR amplicons showed that all isolates detected were members of the D (Dideron) strain of PPV (Spiegel et al., 2004).

In this study, the complete genomes of three plum and one apricot isolates of PPV from Almaty's Pomological Gardens were determined to better characterize the virus isolates and to determine their relationships to each other and to PPV D isolates with sequences deposited in GenBank.

## MATERIALS AND METHODS

### Virus isolates

The PPV isolates from apricot and plum that were subjected to complete genome sequencing were obtained from plants growing in the Pomological Gardens, Almaty, Kazakhstan. Plant samples were prepared and shipped to the CFIA's Sidney Laboratory as described by Spiegel et al. (2004). The isolates used in generating complete genome sequences were A8R2 (from a wild apricot seedling), P5R8 (plum), P7R1 (plum), and P15R8 (plum). Some preliminary screening for the presence of PPV was conducted using a variety of assays. These included the well-validated diagnostic PPV-specific RT-PCR described by Wetzel et al. (1991) that targets the CP region and the RT-PCR assay described by Glasa et al. (2002) that targets the P3-6$K_1$ region.

### Total RNA extractions

Total RNA was extracted from dried plant samples (approximately 100 mg) using the QIAGEN RNeasy Total RNA Plant Mini kit (cat. no. 74904, QIAGEN Canada, Montreal, Canada). The procedure used was essentially that described by the supplier.

### cDNA production, cloning, and sequencing

For Sanger sequencing, cDNA was generated using an approach similar to that described by Myrta et al. (2006). Amplified fragments were gel purified using the QIAGEN MinElute gel extraction kit (cat. no. 28604) and the purified fragments were ligated into the pCR® 2.1-TOPO® vector and cloned using the Invitrogen TOPO® TA cloning kit (cat. no. 45-0641). Sequencing of clones with PPV-derived inserts was carried out at the Nucleic Acid Protein Service Unit, University of British Columbia, Vancouver, Canada, using an Applied Biosystems 3730s 48-capillary DNA Analyzer and POP-7 BigDye Terminator version 3.1 sequencing chemistry.

The complete genome sequence of the plum isolate P15R8 was determined using next-generation sequencing (NGS). Total RNA was extracted using the QIAGEN RNeasy Total RNA Plant Mini kit, but with a modified protocol as described by Kalinowska et al. (2012). Approximately 500 ng total RNA from plum sample P15R8 was submitted to The Centre for Applied Genomics (TCAG), Toronto, Canada. The Illumina TruSeq RNA Library Prep kit was used to enrich for poly(A) RNA, which was used to create a cDNA library. Sequencing was carried out using the Illumina HiSeq 2500.

### Sequence analysis

Traces derived from Sanger sequencing were compiled and sequence analysis carried out using Clone Manager 9 Professional Edition (Scientific and Educational Software, Cary, NC, USA). For NGS-derived sequence data, trimmed reads were mapped to selected PPV D

isolates allowing assembly of the complete genome of isolate P15R8 using CLC Genomics Workbench version 7.5.1 (CLC Bio; http://www.clcbio.com). Phylogenetic analyses were carried out using the neighbour-joining method described by Saitou and Nei (1987). Trees were reconstructed using the NJ tree option of Clustal X version 2 (Thompson et al., 1997). Bootstraps were from 1000 replicates. Analysis was also carried out by the Bayesian method using MrBayes version 3.2.6 (Ronquist et al., 2012). Pairwise distance measurements were carried out using MEGA6 (Tamura et al., 2013).

**RESULTS AND DISCUSSION**

Initially, four plum trees were found positive by RT-PCR, and three of these plum isolates from three different sources were determined to contain a common 6-nt deletion event (Table 1). One of these plum isolates (P5R8), as well as two other plum isolates (P7R1, P15R8) and an apricot isolate (A8R2), were selected for full-genome sequencing. The full-genome sequences of these isolates are deposited in GenBank with accession numbers XA00000, XB00000, XC00000, and XD00000, respectively.

Table 1. *Plum pox virus* isolates detected in plums in Almaty's Pomological Gardens, with three isolates showing a 6-nt deletion in the 5'-terminus region of the coat protein-codding region.

| Cultivar | ID no. | Source | RT-PCR | | Deletion event |
|---|---|---|---|---|---|
| | | | 243-bp amplicon[1] | 836-bp product[2] | |
| Alychagek | P19R18 | Russia | ++[3] | ---[4] | Yes |
| Fidzhenka | P2R15 | Ukraine | ++ | --- | Yes |
| Edinborgs | P3R6 | Kyrgyzstan | +++ | +++ | No |
| Monphor | P5R8 | Kyrgyzstan | +++ | + | Yes |

[1]RT-PCR as described by Wetzel et al. (1991).

[2]RT-PCR as described by Glasa et al. (2002).

[3]Positive results of varying intensities.

[4]Negative results observed.

The wild apricot seedling A8R2 found to be positive for PPV when analysed by sequencing the virus isolate had a genome size of 9786 nt (Table 2). In silico analysis for restriction enzyme sites of the genome of PPV A8R2 revealed a unique site, *Dde*I at position 6577 of the genome. The motifs associated with aphid transmission are present in PPV isolate A8R2, including KITC and PTK motifs present in the deduced HC-Pro region of the virus genome and the DAG motif in the deduced CP region.

Table 2. PPV D isolates from Almaty's Pomological Gardens, Kazakhstan, that had their complete genome sequences determined in this study.

| Isolate | Host | Genome size (nt) | Isolate-specific restriction site |
|---|---|---|---|
| A8R2 | Wild apricot seedling | 9786 | *Dde*I (nt 6577) |
| P5R8 | Plum | 9780[1] | *Bsa*HI (nt 3613) |
| P7R1 | Plum | 9786 | *Hpa*I (nt 8841) |
| P15R8 | Plum | 9786 | *Eco*RII (nt 8688) |

[1]PPV isolate P5R8 has a 6-nt deletion in the 5' terminus of the CP coding region.

The plum-derived PPV isolates P5R8, P7R1 and P15R8 had genome sizes of 9780, 9786 and 9786 nt, respectively. The shorter genome of P5R8 observed reflects the deletion event in the CP coding region, and confirms the deletion event first described by Spiegel et al. (2004). Figure 1 shows an alignment of the deduced amino acid sequences, showing the position of the deletion event isolate in P5R8 (two proline residues) relative to the N

terminus of the coat protein of the virus. The DAG motif of three isolates (A8R2, P5R8 and P7R1) at amino acid positions 11-13 is indicated (Figure 1). All plum isolates sequenced possess the various motifs described above that are associated with aphid transmission. Isolate-specific restriction sites were also identified (Table 2).

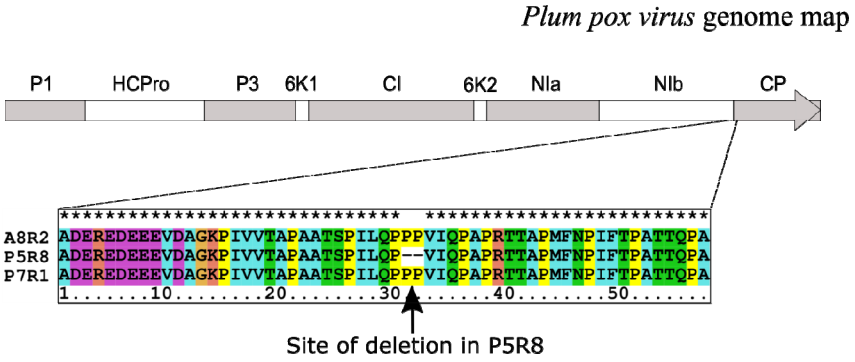


Figure 1. Alignment of the deduced amino acid residues at the N-terminus of the coat protein that encompasses the position of the 2-aa (6-nt) deletion detected in some Kazakhstan PPV isolates.

Phylogenetic analysis based on the complete genome and including the Kazakhstan PPV D isolates with complete genome sequences and representative PPV D isolates from various regions of the world revealed that the Kazakhstan isolates (apricot and plum) were all very closely related (Figure 2), with or without the deletion event. Also, the Kazakhstan PPV isolates are very closely related to the D isolate BIII/2 from Slovakia (Figure 2). These relationships were consistent when either the deduced polyprotein sequence or the nucleotide or deduced amino acid sequence of the RNA-dependent RNA polymerase and CP regions was used. The relationships were always supported by high bootstrap values (Figure 2).

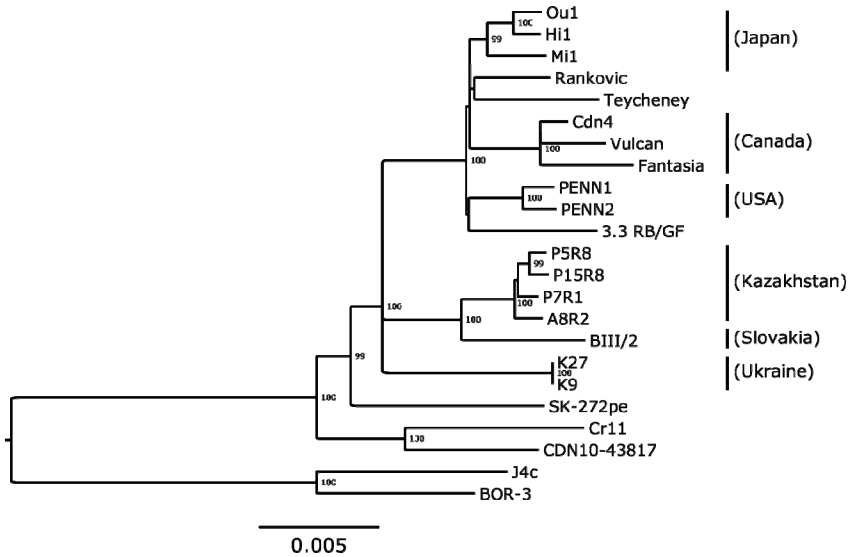


Figure 2. Phylogenetic analysis with trees created using the NJ tree option of Clustal X (Thompson et al., 1997), and based on analysis of some complete genome sequences of PPV D isolates including the four Kazakhstan PPV D isolates with complete genome sequences. Two isolates of strain Rec were included as an outgroup. Percentage bootstrap values (from 1000 replicates) are indicated. Similar relationships were observed with Bayesian analysis (MrBayes version 3.2.6).

Pairwise comparisons further supported the relationships observed in phylogenetic analysis. Percentage identity among the Kazakhstan isolates was 99.8% for both the nucleotide and deduced amino acid sequences of the polyprotein, and 91.1 and 99.2-99.8% for the nucleotide and amino acid sequences of the polyprotein, respectively, when compared with the D isolate BIII/2 from Slovakia (Table 3). This is consistent with the pairwise distances observed and numbers of nucleotides that differ (Table 3). Pairwise comparisons based on the NIb and CP regions gave similar results (data not shown), with identities of 99.7-100% for comparisons among the Kazakhstan PPV D isolates. When compared to the D isolate BIII/2 from Slovakia, the Kazakhstan isolates are approximately 99% identical for the NIb region (nucleotide and amino acid) and over 99% identical for the CP region.

Table 3. Pairwise comparisons of nucleotide sequences (amino acid comparisons in parentheses) showing percentage identity (below the diagonal) and number of nucleotides that differ/total nucleotides and pairwise distances (above the diagonal) of PPV isolates detected in apricot and plums in the Almaty Pomological Gardens, Kazakhstan, and also the PPV D plum isolate BIII/2 from Slovakia. The GenBank accession number for the complete genome sequence of BIII/2 is GU461890.

| Isolate | BIII/2 | A8R2 | P5R8 | P7R1 | P15R8 |
|---|---|---|---|---|---|
| BIII/2 | | 85/9786 $8.6\times10^{-3}$ | 86/9780 $8.7\times10^{-3}$ | 86/9786 $8.7\times10^{-3}$ | 87/9786 $9.0\times10^{-3}$ |
| A8R2 | 99.1 (99.2) | | 23/9780 $2.4\times10^{-3}$ | 23/9786 $2.4\times10^{-3}$ | 24/9786 $2.5\times10^{-3}$ |
| P5R8 | 99.1 (99.3) | 99.8 (99.8) | | 20/9780 $2.0\times10^{-3}$ | 15/9780 $1.5\times10^{-3}$ |
| P7R1 | 99.1 (99.2) | 99.8 (99.8) | 99.8 (99.8) | | 21/9786 $2.2\times10^{-3}$ |
| P15R8 | 99.1 (99.2) | 99.8 (99.8) | 99.8 (99.8) | 99.8 (99.8) | |

Extensive and thorough screening in the past of wild apricot trees growing the Zailiyski Alatou area of the Tien-Shan Mountains has never revealed any PPV-positive wild apricot plants. It would therefore seem that the PPV-positive wild apricot in Almaty's Pomological Gardens was probably the result of transmissions from PPV-infected plum trees. The prevailing winds in the area of the Pomological Gardens are from north to south, and the apricot seedling block found positive for PPV was downwind to the south of plum trees that were found positive for PPV. Aphid transmission of PPV may have been aided by wind movement. All the Kazakhstan PPV D isolates were very closely related to each other (99.8% identity, and common phylogenetic grouping), including both apricot and plum isolates. This suggests a common origin. It is also interesting that the Kazakhstan PPV D isolates always grouped most closely to the PPV D isolate BIII/2 from Slovakia. It is possible that all these isolates may have a common origin, because of their genetic identity.

It may be impossible to determine the exact source of the PPV D isolate detected in the wild apricot seedling in the Pomological Gardens. There is no evidence of PPV-infected wild apricot trees in the Zailiyski Alatou area, the area of origin of the apricot seedlings in the Pomological Gardens. It means that infection probably occurred while the material was maintained in the germplasm collection. This certainly highlights the need for care in establishing and maintaining germplasm collections and in introducing new material into these collections, and the need for rigorous screening of new introductions and perhaps even quarantine of newly collected germplasm before introduction, especially for vector-transmitted pathogens like PPV. This will help prevent the inadvertent loss of valuable, rare and, in some cases, irreplaceable germplasm.

**ACKNOWLEDGEMENTS**

We would like to thank James Phelan for his technical assistance with the assembly of the complete genome of P15R8 using the NGS-derived sequence data. This study was supported in part by funding from the US Agency for International Development (TA-MOU-

01-CA21-01) as part of a collaborative study between scientists from Israel and Kazakhstan, and funding from the Federal Government of Canada under the Plum Pox Monitoring and Management Program (PPMMP) being implemented by the Canadian Food Inspection Agency (CFIA) and Agriculture and Agri-Food Canada (AAFC).

## Literature cited

Glasa, M., Marie-Jeanne, V., Moury, B., Kúdela, O., and Quiot, J.B. (2002). Molecular variability of the P3-6K$_1$ genomic region among geographically and biologically distinct isolates of *Plum pox virus.* Arch. Virol. *147* (*3*), 563–575 http://dx.doi.org/10.1007/s007050200006. PubMed

Kalinowska, E., Chodorska, M., Paduch-Cichal, E., and Mroczkowska, K. (2012). An improved method for RNA isolation from plants using commercial extraction kits. Acta Biochim. Pol. *59* (*3*), 391–393. PubMed

Kostina, K.F. (1969). The use of varietal resources of apricots for breeding. Tr. Gos. Nikitsk. Bot. Sad *40*, 45–63.

Myrta, A., Varga, A., and James, D. (2006). The complete genome sequence of an El Amar isolate of *Plum pox virus* (PPV) and its phylogenetic relationship to other PPV strains. Arch. Virol. *151* (*6*), 1189–1198 http://dx.doi.org/10.1007/s00705-005-0703-x. PubMed

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., and Huelsenbeck, J.P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. Syst. Biol. *61* (*3*), 539–542 http://dx.doi.org/10.1093/sysbio/sys029. PubMed

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. *4* (*4*), 406–425. PubMed

Spiegel, S., Kovalenko, E.M., Varga, A., and James, D. (2004). Detection and partial molecular characterization of two *Plum pox virus* isolates from plum and wild apricot in southeast Kazakhstan. Plant Dis. *88* (*9*), 973–979 http://dx.doi.org/10.1094/PDIS.2004.88.9.973.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. *30* (*12*), 2725–2729 http://dx.doi.org/10.1093/molbev/mst197. PubMed

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. *25* (*24*), 4876–4882 http://dx.doi.org/10.1093/nar/25.24.4876. PubMed

Wetzel, T., Candresse, T., Ravelonandro, M., and Dunez, J. (1991). A polymerase chain reaction assay adapted to plum pox potyvirus detection. J. Virol. Methods *33* (*3*), 355–365 http://dx.doi.org/10.1016/0166-0934(91)90035-X. PubMed