HIGH INTRA-HOST CHERRY VIRUS A POPULATION HETEROGENEITY IN CHERRY TREES IN SLOVAKIA

M. Glasa[1], K. Šoltys[2], Z. Vozárová[1], L. Predajňa[1], N. Sihelská[1], Z. Šubr[1] and T. Candresse[3]

[1]*Institute of Virology, Biomedical Research Centre, Slovak Academy of Sciences, Dúbravská cesta 9, 84505 Bratislava, Slovakia*
[2]*Comenius University Science Park, Comenius University, Ilkovičova 8, 841 04, Bratislava, Slovakia*
[3]*Equipe de Virologie, UMR Biologie du Fruit et Pathologie, INRA, Univ. Bordeaux, CS20032, 33882 Villenave d'Ornon cedex, France*

SUMMARY

Despite the absence of symptoms, next generation sequencing (NGS) of total RNAs from a wild cherry tree revealed a multiple virus infection. Besides prune dwarf virus (PDV) and little cherry virus 2 (LChV-2) the analysis revealed the presence of two genetically divergent cherry virus A (CVA) variants. Their complete genomes could be reconstituted, differing by 14% at the nucleotide level. The presence of these CVA variants in the original cherry tree was confirmed during three consecutive growing seasons by variant-specific RT-PCR. The development of a polyvalent RT-PCR assay targeting the 3' part of the CVA genome allowed the identification of the virus in several other cultivated, wild or ornamental cherries in Slovakia, and the analysis of CVA genomic variability in the amplified region. This has allowed regrouping of known CVA isolates into six major phylogenetic groups, some of which are only recently reported, further extending the known diversity of this virus. In addition, evidence was obtained supporting the recombinant nature of two Slovak isolates and for the presence of "non-cherry" group isolates of CVA in several cherry host plants. Further investigation of the intra-tree CVA variability in five unrelated cherry trees revealed a homogenous lineage in two trees (average nucleotide diversity 0.2-0.5%), while the CVA population in the other trees was composed of variants belonging to three to four different evolutionary lineages (average diversity 7.4-8.8%). The frequent presence of highly divergent molecular variants within the CVA population in single trees further highlights the complex and heterogeneous nature of viral populations infecting perennial hosts.

Keywords: CVA, molecular variants, next-generation sequencing, cherry

Corresponding author: M. Glasa
Fax: + 421.2.54774284
E-mail: Miroslav.Glasa@savba.sk

INTRODUCTION

Cherry virus A (CVA) is a species of the genus *Capillovirus*, family *Betaflexiviridae*. Its single-stranded monopartite RNA genome consists of two overlapping ORFs encoding, respectively, a large polyprotein containing the replication-associated protein translationally fused to the coat protein, and a movement protein (Marais *et al*., 2011; Koinuma *et al*., 2016). Initially described in cherry in Germany (Jelkmann, 1995), CVA is now reported worldwide and has also been found in non-cherry *Prunus* hosts (James and Jelkmann, 1998; Barone *et al*., 2006). Although generally considered as latent in cherries, CVA might still possibly contribute to disorders in various *Prunus* when in mixed infection with other stone fruit viruses (Sabanadzovic *et al*., 2005; Barone *et al*., 2006).

Several studies, mainly based on partial genome characterisation, have highlighted a substantial molecular variability of CVA (Foissac *et al*., 2005; Marais *et al*., 2008, 2012). Very recently, a comparative analysis of CVA sequences generated from NGS data has considerably extended the knowledge on CVA molecular variability (Kesanakurti *et al*., 2017). Besides the high variability, the occurrence of mixed infections and asymptomatic reactions constitutes a substantial challenge in etiology studies and can complicate the accurate diagnosis of the virus.

RNA viruses can evolve as complex viral populations due to the rapid accumulation of mutations (Kutnjak *et al*., 2015). This ability underlines a high adaptative potential, allowing for the rapid selection of biologically distinct variants with a higher fitness in new environments (Garcia-Arenal *et al*., 2001; Safari and Roossinck, 2014). Although many studies have addressed the molecular variability and strain characterisation of plant viruses, less attention has been paid to the analysis of intra-isolate or intra-host diversity of viruses infecting perennial crops (Iglesias *et al*., 2008; Predajňa *et al*., 2012). Yet, the interaction of viral populations in plants generates a number of effects on disease severity or their competition can lead to population replacements (Harper *et al*., 2015).

In this work, complete genome sequences of two CVA variants persistently infecting a single cherry tree were determined, together with partial sequences of additional Slovak CVA isolates from different cultivated, wild or

Table 1. List of CVA isolates characterised in this study and information on the original cherry source.

Isolate	Locality	Original host	Type of plantation	RT-PCR detection using primers		GenBank accession number
				CVA-fw1/ CVA-rev1	CVA_6084F / CVA_6695R	
1046C	Banská Bystrica	*Prunus avium*	wild, old solitary tree	+	+	MF048809, MF048810
BZ2	Bratislava	*P. serotina*	botanical collection	+	+	MF048820 - MF048827
BZ4	Bratislava	*P. serrulata* 'Kiku Shidare Zakura'	botanical collection	+	+	MF048812, MF048836 - MF048843
BZ7	Bratislava	*P. avium*	local genotype, old solitary tree	+	+	MF048813, MF048844 - MF048851
BZ11	Bratislava	*P. serrulata* 'Kanzan'	botanical collection	+	+	MF048852 - MF048859
BB8	Banská Bystrica	*P. avium*	local genotype, old solitary tree	+	+	MF048811, MF048828 - MF048835
DK1	Dolný Kubín	*P. avium*	local genotype, old solitary tree	+	+	MF048814
DK5	Dolný Kubín	*P. avium*	wild, old solitary tree	+	+	MF048816
DK15	Dolný Kubín	*P. avium*	wild, old solitary tree	+	+	MF048815
M2	Most	*P. cerasus*	cultivated, private garden	+	+	MF048817
M3	Most	*P. avium*	cultivated, private garden	+	+	MF048818
M4	Most	*P. avium*	cultivated, private garden	+	+	MF048819

ornamental cherries, showing their high molecular diversity and intra-isolate heterogeneity.

MATERIALS AND METHODS

The CVA isolates analysed here were obtained in 2014-2016 from cultivated, wild and ornamental cherry trees growing in various localities in Slovakia (Table 1).

NGS analysis of 1046C cherry sample and full-length CVA genome determination. Ribosomal RNA-depleted (Ribo-Zero Magnetic kit, Illumina, USA) total RNAs isolated from fresh leaves of a symptomless wild cherry tree (1046C) were used for library preparation and processed with the Illumina transposon-based chemistry (Nextera XT, Illumina, USA). The library was subsequently analysed with 200-bp paired-end sequencing on the Illumina MiSeq platform. *De novo* assembly and mapping against reference and additional complete viral genomes were performed using Geneious 8.1.9 (Biomatters, New Zealand).

The 5'-end of the 1046C CVA genomes were amplified using the 5'/3' RACE Kit, 2nd Generation (Roche, Switzerland) and three specific reverse primers rCVA-610R (5'-ACAGTCAGGTTTGATGGCTC-3'), rCVA-544R (5'-GTTTATGCCACCATTGACTG-3') and rCVA-297R (5'-GGTGCTAGTTGAATGCCACA-3'), following the manufacturer's instructions. The PCR products obtained, of *ca.* 300 bp (overlapping *ca.* 200 bp of the NGS-based sequences) were cloned and 10 clones individually sequenced, enabling the accurate matching of individual sequences to the different CVA variants and thus to completion of the 5' extremity of the sequences.

The NGS-determined 3' end of the CVA genomes were verified by RT-PCR using CVA-7289F (5'-AGGTTAGTTTCTCTTCCCTG-3', sense) and oligod(T) primers.

To independently detect the two CVA variants infecting the 1046C cherry tree, two sets of variant-specific primers were designed and used in RT-PCR, i.e. 1CVA-1320F (5'-CAGTGGGTAGTAGAGCCATC-3', sense)/1CVA-1928R (5'-CCGTTGCTTCTTCAGTTGGC-3', antisense) and 2CVA-1322F (5'-GTAGGCAGTAGGGCCATTAC-3', sense)/2CVA-1905R (5'-GATGATGTTCCCTCCTCGTG-3', antisense). The accuracy of RT-PCR reaction was verified by Sanger-sequencing of cloned or uncloned PCR products, demonstrating the specificity of each primer pair for its cognate variant.

Polyvalent RT-PCR detection and variant analysis. For a general RT-PCR detection of CVA, a two-step RT-PCR protocol was used. cDNA was synthesized from a total RNAs extract obtained from fully developed leaves (NucleoSpin RNA Plant kit, Macherey-Nagel, Germany) using random hexamer primers and AMV reverse transcriptase (both from Promega, USA). An aliquot of cDNA was added to PCR reactions containing TaKaRa Ex Taq polymerase (TaKaRa, Japan). For first PCR, primers CVA-fw1/CVA-rev1 developed by Marais *et al.* (2012) were used. A new primer pair, designed from aligned available CVA genome data (https://www.ncbi.nlm.nih.gov, accessed on May, 2016), was employed to perform additional detection test. The primers CVA_6084F (5'-CCTCCAACTGTGGCWCTCAT-3', sense) and CVA_6695R (5'-TTCTCTATCACTGGTCTCGG-3', antisense) target the 3' part of the CVA genome, spanning nucleotides (nts) 6084-6695 (based on the full-length NC_003689 genome), in the region of overlap between ORF1 and ORF2. For both primer combinations, the following cycling conditions were used: denaturation at 98°C for 1 min, 35 cycles of amplification (98°C for 30 s, 53°C for 30 s and 72°C for 30 s), and a final extension at 72°C for 5 min. The CVA_6084F / CVA_6695R PCR products were either Sanger-sequenced directly or cloned into the pGEM-T Easy cloning vector (Promega, USA) prior to individual sequencing of eight

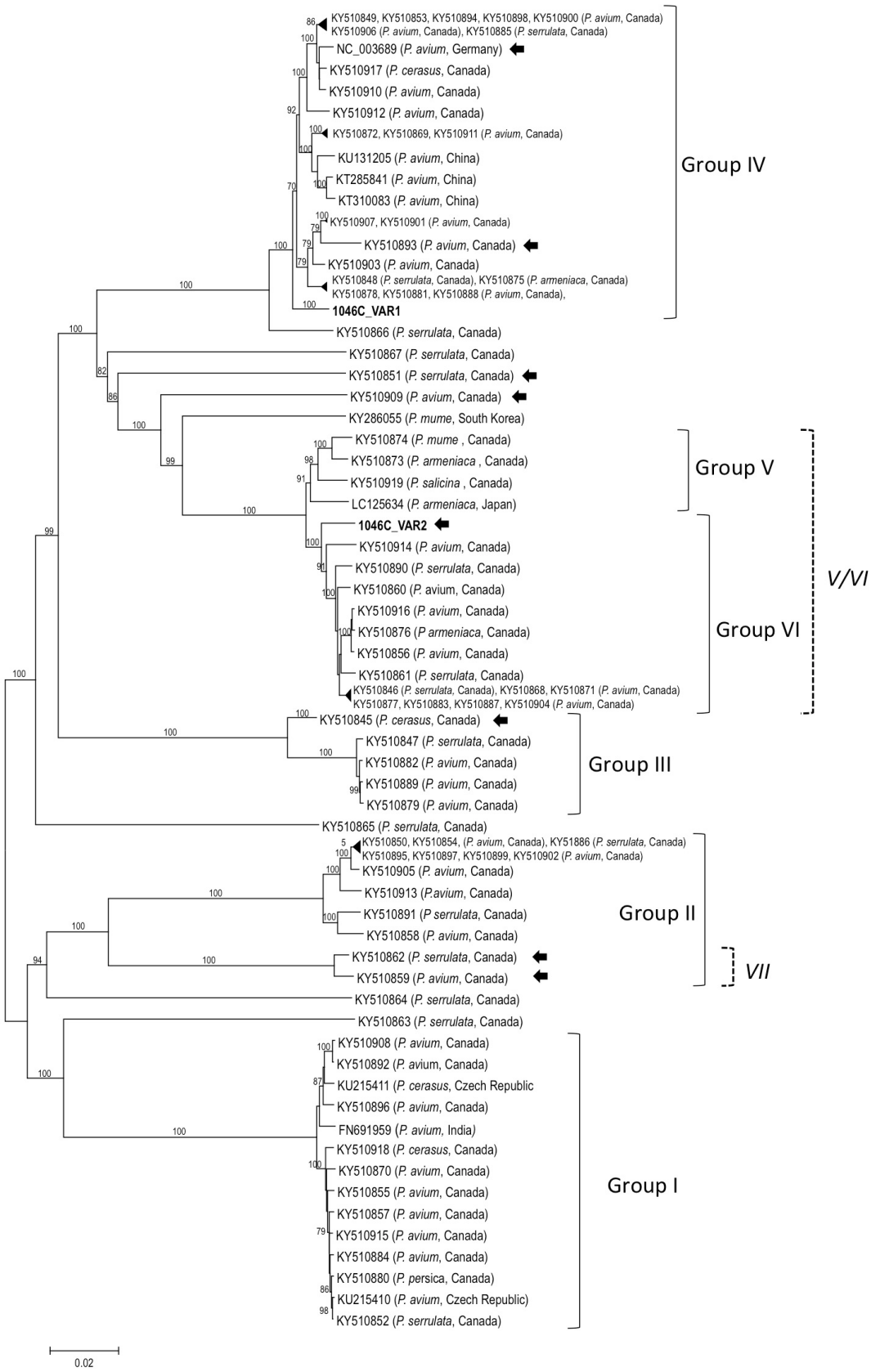

Fig. 1. Phylogenetic tree generated on complete nucleotide genome sequences of CVA isolates. Isolates are identified by their GenBank accession number. The two Slovak variants sequenced in the present study are highlighted in bold. Recombinant isolates identified in this work and by Kesanakurti *et al.* (2017) are marked by an arrow. The phylogenetic groups, based on the work of Kesanakurti e*t al.* (2017) are indicated on the right margin. Alternative grouping proposed in this work is indicated by a dashed parenthesis. The scale bar indicates a genetic distance of 0.02. Bootstrap values higher than 70% (1000 bootstrap re-samplings) are indicated.

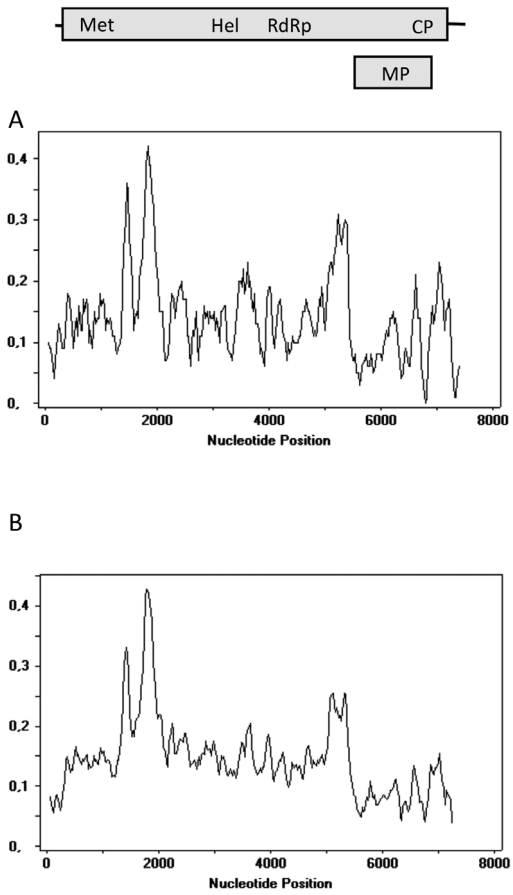

Fig. 2. Nucleotide diversity index calculated along the genomes using DnaSP software with a sliding window of 100 nucleotides moved by steps of 25 nt between A/the aligned sequences of CVA 1046C_VAR1 and VAR2, and B/87 available full-length CVA genomes. Schematic representation of the CVA genome is indicated above the diagrams.

randomly chosen clones from each product. Sequence analyses were performed using either MEGA v.6 (Tamura *et al*., 2013) or DnaSP v.5 (Librado and Rozas, 2009). Searches for potential recombination events employed seven methods, including RDP, GENECONV, Bootscan, Maxchi, Chimaera, SiScan and 3SEQ, implemented in the RDP4 v.4.1 software (Martin *et al*., 2015).

RESULTS

NGS analysis revealed a triple viral species-infection of the 1046C cherry. An apparently symptomless wild cherry tree, labelled as 1046C, growing alone in an urban setting area (GPS coordinates N 48° 44'2.9 E 19° 8'58.7) was previously shown to be infected by prune dwarf virus (PDV) using a RT-PCR assay targeting the viral RNA3 (Predajňa *et al*., 2017). In order to obtain the complete PDV sequence and to reveal the possible presence of other viral agents, high-throughput Illumina sequencing was performed on a library prepared from rRNA-depleted total RNAs isolated from leaves. From 6,914,602 high-quality reads, a total of 1678 contigs were obtained. The BLASTN analysis of these contigs confirmed the presence of PDV and, in addition, revealed the presence of two other viruses, CVA and little cherry virus-2 (LChV-2). Detailed data analysis did not provide evidence for the presence of additional viral agents. For both CVA and LChV-2, this result represents their first identification and characterisation in Slovakia.

Subsequently, mapping of individual reads against the PDV reference genome (NC_008037-39) enabled the nearly complete reconstitution of the sequence of the three genomic RNAs of the Slovak isolate (submitted to Genbank under accession Nos. MF078478, MF078479, MF078480, respectively). In the case of LChV-2, a large contig of *ca*. 5.3 kb was initially obtained and about 78% of the genome was recovered after the reads mapping step (MF078481).

The 1046C cherry tree hosts two divergent CVA variants. Visual inspection of reads mapped against the CVA reference genome (NC_003689) using Geneious default options suggested the presence of distinct CVA variants. Indeed, subsequent blast analysis of contigs revealed the presence of two CVA genomes, closely related to the NC_003689 and LC125634 GenBank sequences, respectively. The repeated mapping against these two genomes using stringent parameters allowed the nearly complete genome reconstitution of two genetically different CVA variants (further referred to as 1046C_VAR1 and 1046C_VAR2). The presence of these two CVA variants and the accuracy of the NGS-derived sequences were unambiguously confirmed by re-sequencing the most variable portions of the genome by variant-specific RT-PCRs, allowing the completion full genome sequences for both isolates, including RACE-based determination of their 5' genome end.

The full-length sequences of the two 1046C variants are 7434 nt (1046C_VAR1, accession No. MF048809) and 7433 nt (1046C_VAR2, accession No. MF048810) long, respectively. The length variability is due to a one nt deletion in the 1046C_VAR2 3' untranslated region (UTR, nt position 7368) as compared to the NC_003689 reference genome. Both genomes have a typical capillovirus organisation, consisting of a 106 nt-long 5' untranslated region (UTR), two large overlapping ORFs of 7026 and 1389 nts and a 301/302 nt-long 3' UTR.

A comparison of the complete genome sequences showed a 13.97% nucleotide sequence divergence between the two variants. The most variable regions include the portions spanning nts 1380-1550, 1700-2000 and 5100-5425, where the divergence exceeds 30%. The same pattern of variability along the genome was noted when all 87 complete sequences, including 85 from GenBank (for accession numbers see Fig. 1) were analysed (Fig. 2).

The divergent nature of the two 1046C variants is also illustrated by the phylogenetic analysis performed on

complete CVA genome sequences. The neighbor-joining tree, reconstructed using strict nucleotide sequence identity distances showed that CVA isolates cluster into at least six major phylogenetic lineages supported by a high bootstrap value, although a few Canadian *P. serrulata* isolates remained ungrouped. As expected, the two 1046C variants clustered in different phylogenetic groups (Fig. 1).

A recombination analysis performed using RDP4 (Martin *et al*., 2015) showed highly supported recombination signals in 1046C_VAR2 sequence and in eight other CVA isolates, with two of them showing evidence of multiple recombination events (for details see Table 3).

Multi-year persistence of the two CVA variants in the 1046C cherry tree. The 1046C cherry tree was periodically sampled twice in the vegetation period (May, August) from 2014 to 2016. To avoid potential irregular virus distribution effect, the samples (composed of 2-3 leaves) were harvested each year from the same branch of the tree, situated in the bottom part of the canopy. The presence of each variant was evaluated using variant-specific RT-PCR assays. Although no effort was made at quantitation, the specific RT-PCR products corresponding to variants 1046C_VAR1 and VAR2 were both obtained at all six time-points (2× in 2014, 2× in 2015 and 2× in 2016). Sanger sequencing of the various PCR products confirmed the specificity of the amplifications and showed no mutations in the amplified genome part of both variants during the three consecutive seasons (data not shown).

Detection of single- and mixed-infections of divergent CVA variants. Eleven additional CVA-infected cherry sources were identified, unequivocally detected by RT-PCR using two different primer combinations, yielding products of *ca*. 280 bp (Marais *et al*., 2012) and *ca*. 610 bp (CVA_6084F/CVA_6695R, see Materials and Methods). The molecular variability of CVA Slovak isolates was assessed by direct Sanger sequencing of CVA_6084F/CVA_6695R-primed PCR products. For nine samples, an unambiguous sequence was generated, allowing assignment of the corresponding isolates to different phylogenetic groups (Group I: DK1, M2, M3; Group II: BZ4; Group V: DK5, DK15, M4). Isolates BB8 and BZ7 clustered in a separate group with previously unassigned Canadian isolates KY510859 and KY510862 (Fig. 3). The partial sequences obtained were not strictly collinear, because the region contained an indel polymorphism site adding or removing up to two aminoacids in the encoded protein.

In the case of samples BZ2 and BZ11, the sequencing electropherograms presented double peaks or unclear reads in a substantial number of positions, indicative of mixed infection of divergent sequences. The amplicons from these two sources, together with three additional ones used as controls (BB8, BZ4 and BZ7), were cloned and individual clones sequenced. As expected, the analysis of the sequences obtained revealed a homogenous lineage

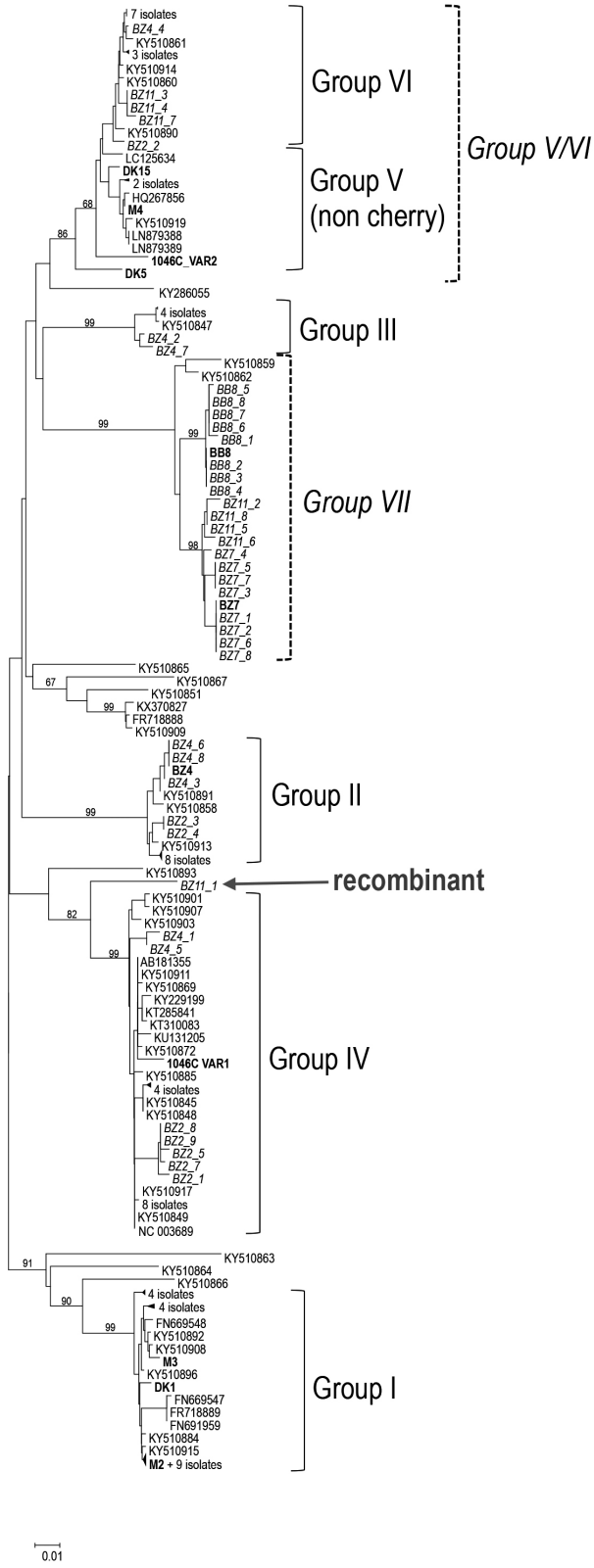

Fig. 3. Neighbor-joining phylogenetic tree generated from partial CVA sequences encompassing the 3' part of the CVA genome (nt 6104-6675). The corresponding sequences of previously characterized isolates are identified by their accession numbers and name. Sequences of Slovak CVA isolates determined by direct sequencing of PCR are in bold and those of individual clones are in italic. Only bootstrap values ≥ 70% (1,000 bootstrap resamplings) are indicated.

Table 2. Analysis of the intra-isolate nucleotide polymorphism in the CVA isolates collected from single cherry trees. After primer removal, the analysed region encompasses *ca*. 570 nucleotides (nt 6104-6675 on the complete NC_003689 genome).

Isolate	Number of analysed sequences/ Number of haplotypes	Haplotype diversity (Hd)	Number of polymorphic sites (S)	Average number of nucleotide differences (k)	Nucleotide diversity (Pi)	Affiliation of haplotypes to the molecular groups (1)
BZ2	8/6	0.929	98	42.107	0.07400	II, IV, VI
BZ4	8/7	0.964	108	50.50	0.08875	II, III, IV, VI
BZ7	8/3	0.679	7	3.286	0.00571	VII
BZ11	8/6	0.929	89	42.571	0.07443	VI, VII, recombinant
BB8	8/4	0.786	4	1.286	0.0022	VII

(1) see Fig. 3

Table 3. The most likely recombination events in the complete sequences of CVA isolates.

Isolate	RDP	GENECONV	BOOTSCAN	MAXCHI	CHIMERA	SISCAN	3SEQ	Breakpoint[1]
1046C_VAR2	2.7×10^{-16}	3.8×10^{-15}	2.3×10^{-16}	–	–	5.5×10^{-3}	–	5431-5649
NC_003689	3.4×10^{-2}	2.4×10^{-2}	1.0×10^{-2}	5.9×10^{-5}	2.4×10^{-5}	3.8×10^{-4}	1.6×10^{-6}	4383-(3745-4454)
KY510909	4.1×10^{-3}	–	5.1×10^{-4}	4.6×10^{-5}	2.6×10^{-6}	1.2×10^{-3}	–	5709-7290
KY510893	5.6×10^{-83}	1.2×10^{-80}	4.8×10^{-84}	1.7×10^{-18}	8.8×10^{-19}	9.1×10^{-21}	6.6×10^{-69}	6365-7362
KY510866	1.3×10^{-13}	7.4×10^{-14}	1.1×10^{-10}	1.1×10^{-16}	2.8×10^{-14}	2.6×10^{-96}	1.0×10^{-59}	64-6002
KY510851	3.2×10^{-15}	1.5×10^{-3}	3.7×10^{-7}	1.7×10^{-9}	4.0×10^{-11}	3.2×10^{-8}	–	5761-7362
KY510845	5.3×10^{-151}	7.5×10^{-148}	6.4×10^{-152}	6.7×10^{-32}	2.6×10^{-32}	7.2×10^{-35}	1.8×10^{-177}	5485-7363
KY510862	2.8×10^{-14}	–	2.2×10^{-13}	4.3×10^{-7}	2.2×10^{-8}	2.8×10^{-7}		4905-5484
	6.7×10^{-6}	–	2.8×10^{-6}	4.8×10^{-2}	1.1×10^{-2}	–	–	5485-6154
	2.0×10^{-4}	3.7×10^{-5}	6.2×10^{-6}	–	–	2.9×10^{-3}	–	7043-7361
KY510859	2.8×10^{-14}	–	2.25×10^{-13}	4.2×10^{-7}	2.2×10^{-8}	2.8×10^{-7}	–	4905-5484
	6.7×10^{-6}	–	2.8×10^{-6}	4.8×10^{-2}	1.1×10^{-2}	–	–	5485-6135
	2.0×10^{-4}	3.7×10^{-5}	6.2×10^{-6}	–	–	2.9×10^{-3}	–	7009-7361

[1]Positions of breakpoints are given on the complete LC125634 sequence.

in the BB8 and BZ7 sources (average nucleotide diversity between sequences 0.2 and 0.5%, respectively). The CVA population in the three other sources was composed of variants belonging to three to four different evolutionary lineages (with an average intra-isolate diversity of 7.4-8.8%, Table 2).

A recombination analysis performed on all sequences obtained for the amplified region indicates that one of the sequences, BZ11-1 generated from a cloned PCR fragment is likely recombinant (Fig. 3). BZ11_1 was detected by 4 of the programs implemented in RDP 4.0, with Bonferroni corrected probabilities between 9.1×10^{-4} and 9.7×10^{-11} (Table 3).

DISCUSSION

Our global knowledge of viral pathogens of plants has significantly increased in the last few years, thanks in part to the application of next-generation sequencing technologies (NGS) (Massart *et al*., 2014). These efforts have often shown the "virome" of perennial plants to be complex (Marais *et al*., 2016; Wu *et al*., 2015; Kesanakurti *et al*., 2017). The old wild-growing cherry tree examined by NGS in the present work has proven to be infected by three viruses (CVA + PDV + LChV-2), although no viral-like symptoms could be recorded during the 2014-2016 evaluation period.

The mixed infections involving two or more isolates or strains of the same virus are frequently detected in perennial crop plants under natural conditions (Kesanakurti *et al*., 2017). However, limited information is available about the persistence and medium/long term evolution of such infections. In the 1046C cherry tree, two divergent variants of CVA were detected during three consecutive seasons, indicating that the two variants are able to coexist over several years without displacement or elimination of one of the isolates. On the contrary, the coexistence of several plum pox virus (PPV) variants in plum trees has been shown to be unstable and only transient (Predajňa *et al*., 2012; Capote *et al*., 2006). No vector of CVA is known up to now (Marais *et al*., 2011), however the presence of two CVA variants in self-rooted and wild cherry trees seems to suggest a possibility of plant-to-plant transmission either through seed or pollen or through as yet unidentified vector(s).

To further investigate the frequency of infections by mixed CVA populations in cherry trees, 11 additional samples were analysed. Based on the direct sequencing of uncloned PCR products, nine of these samples seemed to be infected by a single isolate. However, re-analysis of three samples by sequencing multiple cloned PCR products showed one of them (BZ4) to unexpectedly carry a mixed infection (Table 2, Fig. 3). The failure to detect initially this mixed infection probably reflects a low proportion of one of the variants, that was therefore not identified

when analyzing the sequencing electrophoregram of the uncloned PCR product (Predajňa *et al*., 2012). In total, a minimum of three of the 11 analysed sources proved to be infected by multiple CVA isolates, demonstrating that this situation is rather frequent. These results further highlight the complex and heterogeneous nature of viruses, especially those infecting perennial hosts, and the potential methodological constraints when assessing virus diversity.

CVA has been shown to have a high genetic diversity. Average pairwise nucleotide divergence between isolates in a short fragment of the RNA dependent RNA polymerase reached 9% (Marais *et al*., 2008). The overall mean nucleotide diversity calculated here from all available CVA complete sequences reached 14.4% (±0.2%), further illustrating the substantial level of molecular variability detected within the CVA species (Marais *et al*., 2011). Given this level of diversity, the development of broad-spectrum RT-PCR assay, detecting all virus variants/isolates, can be challenging. In this work, the polyvalent PCR primers CVA_6084F/CVA_6695R were designed, providing an efficient alternative to the CVA-fw1/CVA-rev1v primers (Marais *et al*., 2012), thus increasing the options for the polyvalent RT-PCR detection of CVA targeting a different genome portion, and for the analysis of its genetic diversity.

In this respect, it is noteworthy that two of the six groups identified in the phylogenetic analysis based on the PCR fragment were described only recently by next generation sequencing analysis of Canadian CVA isolates (Kesanakurti *et al*., 2017). Cluster formed by two sequences obtained from the BZ4 tree (BZ4-2 and BZ4-7) belongs to a newly described Group III (Fig. 3). Closely related sequences obtained from BB8 and BZ7 trees cluster with Canadian isolates KY510859 and KY510862 from *P. avium* and *P. serrulata*, respectively, both unassigned in the previous study (Kesanakurti *et al*., 2017). Based on our results, however, these isolates could form an additional phylogenetic group VII (Fig. 3). Taken together these data indicate that a significant proportion of CVA diversity is, at best, still poorly characterized.

The recombination analysis identified two Slovak sequences (BZ11_1 and 1046C-VAR2) likely originating from recombination events. In both cases, one of the recombinant parents was identified in Group V, which interestingly corresponds to "non-cherry" isolates (Marais *et al*., 2012; Kesanakurti *et al*., 2017). The evidence reported here for long term persistence of co-infections involving several CVA variants provides a clear basis for the emergence of recombinant isolates. Although recombination has been identified to shape the molecular diversity of several *Betaflexiviridae* members (e.g. Villamor and Eastwell, 2013; Yoon *et al*., 2014; Marais *et al*., 2015), such a situation has only recently been reported for CVA (Kesanakurti *et al*., 2017). Interestingly, the isolates NC_003689, KY510866, KY510851, KY510859 and KY510862 were not identified as recombinants by Kesanakurti *et al*. (2017). In addition, the two isolates were shown, in the present analysis, to display a pattern involving multiple recombination events and resulting in an unusual mosaic genome structure (Table 3). Our findings thus further reinforce the role of recombination as an important driving force in the evolution of RNA viruses and highlight the need for continuing studies of viral molecular variability, even in the case of well-known and established viruses (Predajňa *et al*., 2017).

Concerning Group V, which so far gathered only non-cherry isolates of CVA, it is noteworthy that some of the sequence variants reported here were obtained from sweet cherry trees, demonstrating that this group should no longer be considered as specific to non-cherry hosts. Moreover, it is evident from the phylogenetic analysis, that groups designed as V and VI by Kesanakurti *et al*. (2017) form a unique monophyletic cluster, which should likely be considered as a single phylogenetic CVA group (shown by a dashed parenthesis in Fig. 1).

NGS has provided a very powerful alternative for detection and identification of plant viruses and viroids without *a priori* knowledge and for discovery of viruses in single or multi-agents disease complexes (Roossinck, 2016; Villamor *et al*., 2016). As shown in this work and elsewhere (Kutnjak *et al*., 2015; Glasa *et al*., 2017), this technology can provide insight into "hidden" intra-isolate heterogeneity of already established viruses. Such information has clear interest for the design of effective and polyvalent detection tools which in turn, as illustrated here, are critical to further our understanding of the variability and of the molecular evolution and epidemiology of plant viruses.

ACKNOWLEDGEMENTS

This work was supported by the grant APVV-0174-12 from the Slovak Research and Development Agency and partially by the grants VEGA2/0036/16 and ITMS313021D075 from the Research & Development Operational Programme funded by the ERDF. The research was conducted within the framework of COST Action FA1407 (DIVAS) supported by COST (European Cooperation in Science and Technology).

REFERENCES

Barone M., Alioto M., Marais A., Candresse T., Ragozzino A., 2006. First report and high prevalence in noncherry hosts of Cherry virus A in Italy. *Plant Disease* **90**: 1459.

Capote N., Gorris M.T., Martínez M.C., Asensio M., Olmos A., Cambra M., 2006. Interference between D and M types of Plum pox virus in Japanese plum assessed by specific monoclonal antibodies and quantitative real-time reverse transcription-polymerase chain reaction. *Phytopathology* **96**: 320-325.

Foissac X., Svanella-Dumas L., Gentit P., Dulucq M.J., Marais A., Candresse T., 2005. Polyvalent degenerate oligonucleotides reverse transcription-polymerase chain reaction: a

polyvalent detection and characterization tool for trichoviruses, capilloviruses, and foveaviruses. *Phytopathology* **95**: 617-625.

Garcia-Arenal F., Fraile A., Malpica J.M., 2001. Variability and genetic structure of plant virus populations. *Annual Review of Phytopathology* **39**: 157-186.

Glasa M., Predajňa L., Šoltys K., Sihelská N., Nagyová A., Wetzel T., Sabanadzovic S., 2017. Analysis of Grapevine rupestris stem pitting-associated virus (GRSPaV) in Slovakia reveals differences in intra-host population diversity and naturally occurring recombination events. *The Plant Pathology Journal* **33**: 34-42.

Harper S.J., Cowell S.J., Dawson W.O., 2015. Finding balance: Virus populations reach equilibrium during the infection process. *Virology* **485**: 205-212.

Iglesias N.G., Gago-Zachert S.P., Robledo G., Costa N., Plata M.I., Vera O., Grau O., Semorile L.C., 2008. Population structure of Citrus tristeza virus from field Argentinean isolates. *Virus Genes* **36**: 199-207.

James D., Jelkmann W., 1998. Detection of Cherry virus A in Canada and Germany. *Acta Horticulturae* **472**: 299-304.

Jelkmann W., 1995. Cherry virus A: cDNA cloning of dsRNA, nucleotide sequence analysis and serology reveal a new plant capillovirus in sweet cherry. *Journal of General Virology* **76**: 2015-2024.

Kesanakurti P., Belton M., Saeed H., Rast H., Boyes I., Rott M., 2017. Comparative analysis of cherry virus A genome sequences assembled from deep sequencing data. *Archives of Virology* **162**: 2821-2828,

Koinuma H., Nijo T., Iwabuchi N., Yoshida T., Keima T., Okano Y., Maejima K., Yamaji Y., Namba S., 2016. First complete genome sequence of Cherry virus A. *Genome Announcements* **9**: 4. pii: e00498-16. doi: 10.1128/genomeA.00498-16.

Kutnjak D., Rupar M., Gutierrez-Aguirre I., Curk T., Kreuze J.F., Ravnikar M., 2015. Deep sequencing of virus-derived small interfering RNAs and RNA from viral particles shows highly similar mutational landscapes of a plant virus population. *Journal of Virology* **89**: 4760-4769.

Librado P., Rozas J., 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451-1452.

Marais A., Candresse T., Svanella-Dumas L., Jelkmann W., 2011. Cherry virus A. In: Hadidi A., Barba M., Candresse T, Jelkmann W. (eds) Virus and Virus-Like Diseases of Pome and Stone Fruits, pp 147-150. APS Press, St. Paul, MN, USA.

Marais A., Faure C., Candresse T., 2016. New insights into Asian Prunus viruses in the light of NGS-based full genome sequencing. *PLoS ONE* **11**: e0146420. doi:10.1371/journal.pone.0146420

Marais A., Faure C., Mustafayev E., Candresse T., 2015. Characterization of new isolates of Apricot vein clearing-associated virus and of a new Prunus-infecting virus: Evidence for recombination as a driving force in Betaflexiviridae evolution. *PLoS ONE* **10**: e0129469.

Marais A., Svanella-Dumas L., Barone M., Gentit P., Faure C., Charlot G., Ragozzino A., Candresse T., 2012. Development of a polyvalent RT-PCR detection assay covering the genetic diversity of Cherry capillovirus A. *Plant Pathology* **61**: 195-204.

Marais A., Svanella-Dumas L., Barone M., Ragozzino A., Gentit P., Candresse T., 2008. Genomic diversity of Cherry capillovirus A (CVA) and suitability of various assays for its detection. *Acta Horticulturae* **781**: 37-45.

Martin D.P., Murrell B., Golden M., Khoosal A., Muhire B., 2015. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evolution* **1**: vev003, https://doi.org/10.1093/ve/vev003

Massart S., Olmos A., Jijakli H., Candresse T., 2014. Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus Research* **188**: 90-96.

Predajňa L., Šubr Z., Candresse T., Glasa M., 2012. Evaluation of the genetic diversity of Plum pox virus in a single plum tree. *Virus Research* **167**: 112-117.

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. *European Journal of Plant Pathology* **147**: 877-885.

Roossinck M.J., 2017. Deep sequencing for discovery and evolutionary analysis of plant viruses. *Virus Research* **239**: 82-86.

Sabanadzovic S., Abou-Ghanem-Sabanadzovic N., Rowhani A., Grant J.A., Uyemoto J.K., 2005. Detection of Cherry virus A, Cherry necrotic rusty mottle virus and Little cherry virus 1 in California orchards. *Journal of Plant Pathology* **87**: 173-177.

Safari M., Roossinck M.J., 2014. How does the genome structure and lifestyle of a virus affect its population variation? *Current Opinion in Virology* **9**: 39-44.

Tamura K., Stecher G., Peterson D., Filipski A., Kumar S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**: 2725-2729.

Villamor D.E., Mekuria T.A., Pillai S.S., Eastwell K.C., 2016. High-throughput sequencing identifies novel viruses in nectarine: insights to the etiology of stem-pitting disease. *Phytopathology* **106**: 519-527.

Villamor D.E.V., Eastwell K.C., 2013. Viruses associated with rusty mottle and twisted leaf diseases of sweet cherry are distinct species. *Phytopathology* **103**: 1287-1295.

Wu Q., Ding S.W., Zhang Y., Zhu S., 2015. Identification of viruses and viroids by next-generation sequencing and homology dependent and homology independent algorithms. *Annual Review of Phytopathology* **53**: 425-444.

Yoon J.Y., Joa J.H., Choi K.S., Do K.S., Lim H.C., Chung B.M., 2014. Genetic diversity of a natural population of Apple stem pitting virus isolated from apple in Korea. *The Plant Pathology Journal* **30**: 195-199.

Received July 19, 2017
Accepted August 30, 2017