Virus Research 181 (2014) 1–5

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Short communication

Deep-sequencing analysis of an apricot tree with vein clearing symptoms reveals the presence of a novel betaflexivirus

Toufic Elbeaino[a,*], Annalisa Giampetruzzi[b], Angelo De Stradis[b], Michele Digiaro[a]

[a] Istituto Agronomico Mediterraneo di Bari, Via Ceglie 9, 70010 Valenzano, Bari, Italy
[b] Istituto di Virologia Vegetale del CNR, sezione di Bari, Via Amendola 165/A, 70126 Bari, Italy

ARTICLE INFO

Article history:
Received 4 October 2013
Received in revised form
20 December 2013
Accepted 21 December 2013
Available online 31 December 2013

Keywords:
Apricot
Betaflexiviridae
Deep-sequencing
Apricot vein clearing-associated virus
Genome organization
Phylogeny

ABSTRACT

Deep-sequencing technology applied on double stranded RNA recovered from an apricot tree with vein clearing symptoms allowed the identification of a novel virus with a single-stranded RNA genome, for which the provisional name apricot vein clearing-associated virus (AVCaV) is proposed. Its genome comprises 7315 nt, excluding the poly(A) tail, covering four open reading frames (ORFs). The putative virus-encoded proteins, *i.e.*, replicase (REP), movement protein (MP), coat protein (CP) and nucleic acid-binding protein (NB), had an estimated molecular weight of 192.5, 32.15, 25.5 and 16.1 kDa, respectively and shared the highest identity (*ca.* 40%) with citrus leaf blotch virus (CLBV) and with orthologs of other known members of the family *Betaflexiviridae*. The phylogenetic trees constructed with the sequences of the entire replication-associated polyproteins and the putative CP showed incongruent allocations of AVCaV within the genus *Citrivirus* or as an outgroup species close to the genus *Vitivirus*, respectively. The peculiar organization of its genome (four ORFs), different from that typical of members of *Citrivirus* (three ORFs) and *Vitivirus* (five ORFs) genera, makes likely AVCaV a novel member of an unassigned genus of the family *Betaflexiviridae*. In RT-PCR assays, AVCaV was found to infect only one out of 39 varieties of apricot tested; thus, suggesting to be limitedly spread.

Until a few years ago, only a limited number of viruses infecting apricot had been known, which included ilarviruses [Prunus necrotic ringspot virus (PNRSV), apple mosaic virus (ApMV), prune dwarf virus (PDV), Tulare apple mosaic virus (TAMV)], nepoviruses [Arabis mosaic virus (ArMV), tomato ringspot virus (ToRSV)], a sadwavirus [strawberry latent ringspot virus (SLRV)], and a potyvirus [plum pox virus (PPV)]. These viruses are easily transmissible onto herbaceous hosts.

The use of molecular techniques has increased significantly the list of viruses infecting apricot, especially species belonging to the family *Betaflexiviridae*, *i.e.*, capilloviruses (cherry virus-A, CVA), foveaviruses (apricot latent virus, ApLV), trichoviruses (apple chlorotic leaf spot virus, ACLSV; apricot pseudo chlorotic leaf spot virus, APsCLSV; cherry mottle leaf virus, CMLV; peach mosaic virus, PcMV), and unassigned viruses (cherry green ring mottle virus, CGRMV and cherry necrotic rusty mottle virus, CNRMV) (Adams et al., 2011; Martelli et al., 2011).

Members of the family *Betaflexiviridae*, order *Tymovirales*, are flexuous filaments, usually 12–13 nm in diameter and from 600 to over 1000 nm in length, containing a single molecule of linear ssRNA of about 5.9–9.0 kb. Depending upon the genus, the polyprotein can contain from three to six genes. The ORF1-encoded product (190–250 kDa) contains the conserved domains for methyltransferase (Mtr), helicase (Hel) and RNA dependent RNA polymerase (RdRp) activity. Smaller ORFs encode the proteins involved in cell-to-cell movement, either a single movement protein (MP) of the "30K" superfamily (*Capillovirus*, *Citrivirus*, *Trichovirus*, *Vitivirus*) or a "triple gene block" (TGB). The coat protein (CP) gene always follows the MP(s), and in some genera (*Carlavirus*, *Vitivirus*, and some trichoviruses) a final ORF encodes a protein with a zinc binding finger motif. Nowadays, six genera (*Capillovirus, Carlavirus, Citrivirus, Foveavirus, Trichovirus*, and *Vitivirus*) are included in the family *Betaflexiviridae*, together with a certain number of species that are still unassigned (Adams et al., 2011).

During a field survey in apricot orchards, vein clearing and mottling symptoms were observed on young leaves of a 3-year-old apricot (*Prunus armeniaca* L.) tree of unknown cultivar, denoted VC1 (Fig. 1a). Symptomatic leaves were collected and subjected to laboratory analysis to define the aetiology. In preliminary assays conducted by ELISA and/or RT-PCR, none of the main known apricot-infecting viruses [*i.e.*, PNRSV, ApMV, ACLSV, PPV, PDV, ALRSV, SLRV, ToRSV, ApLV] was detected, except for Plum bark

* Corresponding author. Tel.: +39 0804606352; fax: +39 0804606503.
E-mail address: elbeaino@iamb.it (T. Elbeaino).

0168-1702/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.virusres.2013.12.030

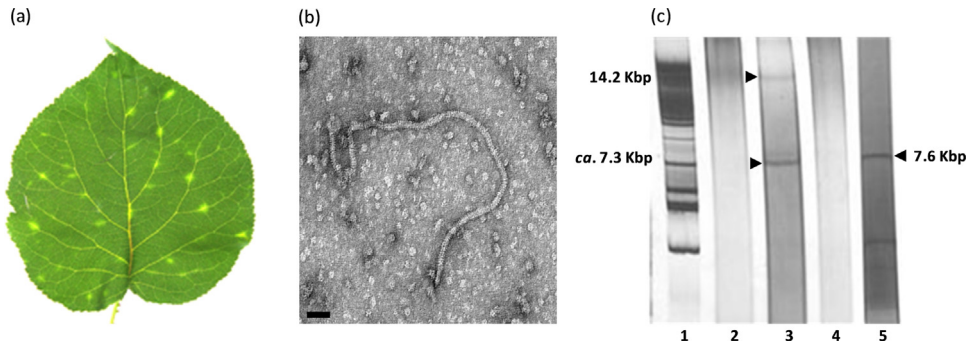

Fig. 1. (a) Symptoms of vein clearing and blotching observed on a leaf of apricot (accession VC1) (b) *Flexiviridae*-like virion in a leaf dip from a symptomatic leaf of accession VC1. The bar represents 50 nm. (c) Electropherogram of dsRNA bands extracted from accession VC1. Lane 1: DNA molecular marker (λ/*Hind*III). Lane 2: healthy grapevine seedling with no dsRNA band. Lane 3: dsRNA pattern extracted from young symptomatic leaves of apricot accession VC1, with the highest (14.2 kbp) and the lowest (7.3 kbp) bands corresponding to PBNSPaV and AVCaV, respectively. Lane 4: healthy apricot plant with no dsRNA band. Lane 5: dsRNA band (7.6 kbp) of grapevine virus A (GVA) extracted from grapevine.

necrosis stem pitting associated virus (PBNSPaV) (Abou Ghanem-Sabanadzovic et al., 2001). PBNSPaV infects apricots and other *Prunus* species (almond, peach, plum, cherry) inducing symptoms limited to the trunk (Di Terlizzi and Savino, 1995; Amenduni et al., 2004; Amenduni et al., 2005).

The presence of possible mechanically transmissible viruses in VC1 was investigated by inoculating sap from young symptomatic leaves ground in 0.1 M phosphate buffer and 2.5% nicotine on a range of herbaceous hosts including species of Solanaceae (*Nicotiana benthamiana*, *N. occidentalis*, *N. tabacum* cv. Samsun), Chenopodiaceae (*Chenopodium quinoa*, *C. amaranticolor*), Amaranthaceae (*Gomphrena globosa*), and Fabaceae (*Vigna unguiculata*, *Phaseolus vulgaris* cv. La Victoire) grown in a glasshouse at 24 °C. All hosts remained symptomless, and no virus was transmitted from any of them by sub-inoculations to new sets of healthy plants.

Leaf dips observed with a Philips Morgagni 282D electron microscope disclosed the presence of two different types of filamentous flexuous particles, the longest of which were *ca.* 1500 × 13 nm in size and were decorated by antibodies to PBNSPaV (not shown). The smaller particles were 950–1100 × 13 nm, exhibited surface patterns with visible cross-banding (Fig. 1b) resembling virions of *Flexiviridae* (Martelli et al., 2007) and were not decorated by the PBNSPaV antiserum. These particles were; therefore, tentatively identified as the putative agent of the vein clearing symptoms to which the provisional name of Apricot vein clearing-associated virus (AVCaV) was assigned.

Leaf veins (15–30 g) of VC1 were used to recover dsRNAs by phenol/chloroform extraction and chromatography through cellulose CF-11 column in the presence of 17% ethanol followed by chromatography through cellulose CC41 (Whatman, USA) (Elbeaino et al., 2009) and selective enzymatic digestion (DNase and RNase) (Saldarelli et al., 1994). Extracts were then passed through Micro Bio-Spin 30 Columns in RNase-free Tris buffer (Bio-Rad Laboratories, Hercules, CA, USA), analyzed in 6% PAGE and silver-stained. Electropherograms showed a dsRNA pattern characterized by the presence of two bands *ca.* 14 and 7.5 kb in size (Fig. 1c), which were attributed to PBNSPaV and to AVCaV, respectively.

To determine the genome sequence of AVCaV a next generation sequencing approach was followed using Illumina technology. A cDNA sequencing library was synthesize from purified viral dsRNAs using a TruSeq RNA Sample Prep Kits v2 (Illumina) and sequenced in a run of 50 bp-single read at University of Bari, Department of Soil, Plants and Food Sciences (http://www.selge.uniba.it) on a HiScan SQ apparatus (Illumina, San Diego, USA). The raw reads were filtered for quality, reduced to unique reads and then were *de novo* assembled into larger contigs using the Velvet Software 1.2.08 (Zerbino and Birney, 2008) with a kmer of 23. Assembled contigs were then screened for sequence homologies using BLASTX (cutoff e-value 10^{-6}) and BLASTP (http://www.ncbi.nlm.nih.gov/). Genomic RNA was reconstructed by manually aligning homologous contigs on the most closely related sequences found in the GenBank database and using the CLC software (CLCbio, Aarhus, Denmark). Once assembled, the viral genome was re-sequenced through several RT-PCR amplifications, using a battery of sense and antisense specific primers designed on deep sequencing-generated reads, to ensure at least fivefold nucleotide coverage. The 5′ and 3′ terminal sequences were determined using 5′/3′ RACE PCR (Roche Diagnostic, Milan, Italy). All amplified products were ligated in StrataClone™ PCR Cloning vector pSC-A (Stratagene, USA), cloned in *Escherichia coli* DH5α or SoloPACK cells, and custom sequenced (Primm, Milano, Italy). Nucleotide and protein sequences were analyzed with the assistance of the DNA Strider 1.1 program (Marck, 1988). Multiple alignments of nucleotide and amino acid sequences were obtained using the default options of CLUSTALX 1.8 (Thompson et al., 1997). Search for homologies with proteins from the Protein Information Resources database (PIR, release 47.0) was done with the FASTA (Pearson and Lipman, 1988), BlastX and BlastP (Altschul et al., 1997) programs. Conserved domains in putative products of the virus genome were identified using CDD (Marchler-Bauer et al., 2011), Pfam (Finn et al., 2010), and TMHMM (Sonnhammer et al., 1998). Secondary structures were predicted with CyloFold software (Bindewald et al., 2010).

For phylogenetic analysis, amino acid sequences were aligned using "MUSCLE" (Edgar, 2004), and trees were generated by the "Maximum likelihood" method (Jones et al., 1992), applying the JTT matrix and pairwise gap deletion options implemented in MEGA5.05 (Tamura et al., 2011). Branches with less than 50% bootstrap support were collapsed.

Deep sequencing showed a total of 8.151.941 unique reads. After adapters removal, size selection (21–51 nt) and filtering for non-viral RNAs, a data set of 536,215 short reads were obtained and subjected to *de novo* assembly, generating a total of 252,188 contigs, ranging in size from 45 to 3681 nt with a N50 of 133 nt. BLAST search of the resultant contigs in the NCBI virus database, using a cutoff e-value of 10^{-6} for BLASTX, identified 74 contigs raging from 45 to 436 nt in size with an average of 105.5 nt, covering 96.3% (6802 nt out of 7062) of AVCaV genome, whereas 273 contigs, ranging from 45 to 286 nt, covered 92.6% of the PBNSPaV genome. In Blast analysis, no homology with other known viruses has emerged from the deep sequencing generated-clones. The complete genome sequence of the PBNSPaV isolate was assembled (14,208 nt) and showed an identity of *ca.* 98.5% with the sequence of the isolate retrieved from Genbank (accession number EF546442).