Japanese persimmon (*Diospyros kaki* Thunb.) originated in eastern Asia, and has been grown for centuries. This plant has been cultivated commercially mainly in China, Korea, and Japan, and recently also in Brazil, Italy, Israel, and other countries. Some graft-transmissible disorders of possible viral origin have been reported (Kusano & Shimomura, 1994; Scott & Payne, 1988; Yanase et al., 1992). For example, a graft-transmissible fruit apex disorder that caused necrosis of sieve tissue in the fruit of cultivars Hiratanenashi and Tonewase (a bud mutation of Hiratanenashi) was reported to harbour filamentous virus-like particles, based on transmission electron microscopy (Imada, 2000). Recently, three apscaviroids, citrus viroid VI (CVd-VI, synonym: CVd-OS), apple fruit crinkle viroid (AFCVd) and persimmon viroid (Pvd), were identified from Japanese persimmons. One of the complete genomes consisted of 13,467 nt and encoded six genes similar to those of plant rhabdoviruses. The virus formed a distinct cluster in the genus *Cytorhabdovirus* with lettuce necrotic yellows virus, lettuce yellow mottle virus and strawberry crinkle virus in a phylogenetic tree based on the L protein (RNA-dependent RNA polymerase, RdRp). The other consisted of 7,475 nt and shared a genome organization similar to those of some insect and fungal viruses having dsRNA genomes. In a phylogenetic tree using the RdRp sequence of several unassigned dsRNA viruses, the virus formed a possible new genus cluster with two insect viruses, *Circulifer tenellus* virus 1 and *Spissistilus festinus* virus 1, and one plant virus, cucurbit yellows-associated virus.

Metagenomic deep-sequencing analyses have enabled us to identify unknown viral molecules that have a detectable sequence homology to known viruses without our prior knowledge of their identity or laborious cloning work (Al Rwahni et al., 2009; Kreuze et al., 2009; Wu et al., 2010). To apply deep-sequencing techniques to Japanese persimmon, nucleic acids were extracted from two Japanese persimmon trees exhibiting fruit apex disorder in some fruits. Extracts were obtained from Kaki13-14 and Kaki13-7 (Table 1), comprising 1 g leaf petioles and midveins from Kaki13-14 using the RNA isolation method described by Reid et al. (2006), and a mixture of 0.05 g leaf petioles and midveins from Kaki13-14 and 0.05 g leaf petioles from Kaki13-7 using a Plant/Fungi Total RNA Purification kit (Norgen). A mixture of the two extracts was sent to Hokkaido System Science Co. for processing and sequencing on an Illumina Genome Analyser II as single-end 50 bp reads. Briefly, a 10 μg mixture was treated with S1 nuclease (Takara) to enrich for dsRNA. The preparation was electrophoresed in a 1% low-melting-point agarose gel, and a gel slice containing an approximately 10 kb band was excised and purified. The purified extract (120 ng) was processed and sequenced using an Illumina mRNA-Seq Sample Prep kit, an Illumina Standard Cluster Generation kit v4, and an Illumina Sequencing kit v4. Deep-sequencing analysis of the fragments yielded approximately 1106 Mb containing 21,478,090 reads once the adaptor sequences were trimmed. Assembly of the sequences using the Velvet program (Zerbino & Birney, 2008) with 31 or 27 nt as the minimal overlapping lengths (k-mer) yielded 27,047 contigs of 61–4326 nt and 156,399 contigs of 53–2089 nt,
respectively. Searches using BLASTN and BLASTX (http://blast.ncbi.nlm.nih.gov/Blast.cgi) detected two clusters of contigs having detectable sequence homology to known viral RNA sequences.

One cluster included 13 contigs of 84–598 nt showing the lowest e-values with the L [RNA-dependent RNA polymerase (RdRp)], glyco- (G) and nucleocapsid (N) proteins of several rhabdoviruses (Dietzgen et al., 2012). For subsequent analyses, initial primers were designed based on the nucleotide sequences of four of the contigs at the determined positions of nt 786–869, 5020–5477, 8429–8567 and 10 584–10 768. The other cluster included eight contigs (53–61 nt) showing the lowest e-values with the RdRp of two insect viruses, *Circulifer tenellus* virus 1 (CiTV1) and *Spissistilus festinus* virus 1 (SpFV1), and one plant virus, *cucurbit yellow-associated virus* (CuYAV), with dsRNA genomes (Coffin and Coutts, 1995; Spear et al. 2010). Initial primers were designed based on the nucleotide sequences of two contigs at the determined positions of nt 5544–5596 and 6191–6243.

For Sanger sequencing, nucleic acids were extracted by the method of Nakaune & Nakano (2006) or using a Plant/Fungi Total RNA Purification kit. The reverse transcription (RT)-PCR with specific primers, the molecular agent, was performed by the method of Nakaune & Nakano (2006) or using LATaq reverse transcriptase (Clontech) with a mixture of random and oligo(dT) primers. The PCR was carried out using the protocol of Nakaune & Nakano (2006) or using LA Taq (Takara) to fill sequence gaps. 5′-RACE and 3′-RACE were also repeatedly performed until the 5′- and 3′-terminal ends were fully analysed, using a SMARTer RACE cDNA Amplification kit (Clontech) with primers based on the derived nucleotide sequences. A poly(A) tail was added to the RNA using *Escherichia coli* poly(A) polymerase (New England Biolabs) in advance of the RACE in which oligo(dT) was used for primer annealing. RACE-PCR bands were purified using a QIAEX II kit (Qiagen) and cloned using a TA cloning kit (Invitrogen), and more than three clones were obtained following the method of Shimizu et al. (2011), modified to use LA Taq in the colony PCR. Single PCR bands were sequenced directly in both directions after removal of the primers, using an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) with a BigDye Terminator v3.1 Cycle Sequencing kit (Applied Biosystems) using a primer-walking strategy. The nucleotide sequences were analysed using GENETYX v9 (GENETYX). Phylogenetic trees were constructed using the neighbour-joining algorithm in CLUSTAL W (Thompson et al., 1994).

The one genome completely determined by Sanger sequencing consisted of 13 467 nt. The genome encoded six ORFs, five of which showed the lowest e-values in the BLASTP analysis with the N, phosphor- (P), matrix (M), G and L proteins of some plant-adapted rhabdoviruses (Dietzgen et al., 2012), with which they had analogous genomic positions (Fig. 1a). Subsequent analyses showed that 39 (93–400 nt) of 27 047 contigs in a 31 k-mer assembly (0.14 %) and 68 (53–598 nt) of 156 399 contigs in a 27 k-mer (0.04 %) were derived from the genome. To check the graft transmissibility of the viral molecule, scions from Kaki13-14 were grafted to four ‘Fuyu’ seedlings. Reverse transcription (RT)-PCR with specific primers, the homologous 5RhF1 (5′-GGATGGCCTTCTGATGTTGG-3′, nt 7575–7594) and the complementary 5RhC1 (5′-TGATGATTGCTATGGCGAT-3′, nt 7852–7871), were carried out using a OneStep RT-PCR kit (Qiagen) as described by Shimizu et al. (2011) with some modifications. At 18 months post-inoculation, RT-PCR detected the molecule in all four grafted seedlings that had previously tested negative. In addition, the molecule was detected from the Kaki13-7 and Kaki13-14 trees, as well as the Kaki13-15 trees clearly exhibiting fruit apex disorder (Table 1). However, the molecule was also detected from asymptomatic Kaki13-2 and Kaki13-12 trees (Table 1); thus the amplified viral sequence was not always associated with disease symptoms, showing that further investigation of the causes of fruit apex disorder is required. Some of the RT-PCR bands were sequenced directly to confirm their specific amplification.

The viral agent, provisionally named persimmon virus A (PeVA), clustered with lettuce necrotic yellow eyes virus (LNYV), lettuce yellow mottle virus (LYMoV) and strawberry crinkle virus (SCV) in the genus *Cytorhabdovirus* in a phylogenetic tree of the conserved L protein of plant-adapted rhabdoviruses (Fig. 2a). The amino acid sequence identities were 53, 42 and 43 % with SCV, LNYV and LYMVo, respectively. Other amino acid sequence identities with LNYV and LYMVo of 30–32 % for the N protein, 15–17 % for the P protein, 13–16 % for the M protein and 25–26 % for the G protein were the highest among those with plant rhabdoviruses whose complete genomes were available.

### Table 1. Multiple comparison test of means of numbers of fruits showing apex disorder among 14 fruits from Japanese persimmon trees of cultivars Hiratanenashi and Tonewase infected with persimmon virus A (PeVA) and persimmon latent virus (PeLV)

<table>
<thead>
<tr>
<th>Tree</th>
<th>PeVA</th>
<th>PeLV</th>
<th>Grafting*</th>
<th>Mean ± SEM†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kaki13-2</td>
<td>+</td>
<td>+</td>
<td>1998 (T)</td>
<td>0.00 ± 0.00 a</td>
</tr>
<tr>
<td>Kaki13-12</td>
<td>+</td>
<td>+</td>
<td>1998 (T), 2006 (H)</td>
<td>0.00 ± 0.00 a</td>
</tr>
<tr>
<td>Kaki13-13</td>
<td>+</td>
<td>+</td>
<td>1998 (T)</td>
<td>0.25 ± 0.13 a</td>
</tr>
<tr>
<td>Kaki13-14</td>
<td>+</td>
<td>+</td>
<td>1998 (T), 2006 (H)</td>
<td>2.50 ± 0.64 a</td>
</tr>
<tr>
<td>Kaki13-7</td>
<td>+</td>
<td>+</td>
<td>2003 (T)</td>
<td>1.75 ± 0.59 a</td>
</tr>
<tr>
<td>Kaki13-11</td>
<td>+</td>
<td>+</td>
<td>2002 (T), 2006 (T)</td>
<td>3.25 ± 0.55 ab</td>
</tr>
<tr>
<td>Kaki13-15</td>
<td>+</td>
<td>+</td>
<td>2002 (T), 2006 (T)</td>
<td>7.25 ± 1.01 b</td>
</tr>
</tbody>
</table>

*+, Detected; −, not detected.

*Years of grafting of cultivars Hiratanenashi (H) and Tonewase (T, a bud mutation of Hiratanenashi). Note that tree ages, cultivars and origins of scions were not necessarily the same as each other.

†Means of numbers of fruits showing apex disorder among 14 fruits randomly selected each for 4 years (2009–2012). Means followed by the same letter in a column are not significantly different according to Tukey’s test at $P<0.05$ (n=14).
in the databases. Interestingly, BLASTP analysis of the ORF3 protein of PeVA showed no significant identity to rhabdovirus proteins, but had the lowest e-value (0.006) with the movement protein (MP) of the DNA-B components of the bipartite begomovirus kudzu mosaic virus (GenBank accession no. ACL78778). In contrast, the ORF3 (4b) proteins of LNYV and LYMoV are putative MPs showing similarities especially with MPs of capilloviruses and trichoviruses (Dietzgen et al., 2006; Heim et al., 2008). These viruses appeared to have shared their MPs by horizontal gene transfer or convergent evolution. Twenty-six of the 34 terminal nucleotides of the 3′ leader and 5′ trailer sequences of PeVA were complementary (Fig. 1b), and possibly could form a panhandle structure reported for some rhabdoviruses (Heim et al., 2008). A 1 nt overhang in the trailer sequence of PeVA was observed, although its significance is unknown (Fig. 1b). Furthermore, a characteristic consensus sequence was found between each ORF, although some differential nucleotides were observed, especially in the region between the 3′ leader and the N protein (Fig. 1c). A part of the consensus sequence, 3′-AUUAUU-UU-5′, was well conserved among intergenic junction sequences of other rhabdoviruses (Heim et al., 2008).

PeVA apparently has genetic characteristics of plant rhabdoviruses, although its other properties have not been investigated. The species demarcation is primarily based on host range and vector specificity (Dietzgen et al., 2012). At the least, PeVA is the first putative rhabdovirus known to be infectious to persimmons. The sequence identity data may be insufficient for accurate species discrimination due to the low number of completely sequenced genomes in the genus (Dietzgen et al., 2012). However, 41–65% amino acid identities between serologically unrelated LNYV and LYMoV appeared to be sufficiently large to regard them as distinct cytorhabdovirus species (Heim et al., 2008).
particles by trial electron microscopy analyses of negative-
ly stained leaf sap from Kaki13-14. Electron microscopy
analysis of purified virions and investigation of the bio-
lological properties of PeVA such as host range and vector
specificity should be further performed.

The other genome completely determined by Sanger se-
quencing consisted of 7475 nt. Subsequent analyses showed
that 379 (61–88 nt) of 27 047 contigs in the 31
k-mer
assembly (1.4 %) and 931 (53–74 nt) of 156 399 contigs in
the 27
k-mer (0.6 %) were derived from the genome. RT-
PCR with specific primers, the homologous Chry-f (5
9
- CGATCCACTGACCTGATCAAC-3’, nt 5964–5984) and
the complementary 272568r (5’-TAGAGCAGCGCAAAT-
ACTC-3’ nt 6195–6214), detected the molecule in all four
seedlings that were grafted with Kaki13-14. However, the
molecule was not detected from the Kaki13-15 tree that
clearly exhibited fruit apex disorder symptoms but was
detected in the asymptomatic Kaki13-2 and Kaki13-12 trees
(Table 1), demonstrating unequivocally that this second
virus was not causally associated with fruit apex disorder.

The viral agent, provisionally named persimmon latent
virus (PeLV), encoded two major ORFs, each of which
showed significant
-e-values with the PArp of CiTV1 and
SpFV1, and RdRp of CiTV1, SpFV1 and CuYAV in the
BLASTP search (Fig. 1d). Spear
et al. (2010) proposed that
CiTV1 and SpFV1, tentatively with CuYAV for which only
limited information was available, be included in a new
genus. Criteria defining the genus may include a non-seg-
mented dsRNA genome, lack of genome encapsidation in a
conventional virion, a gene complement including a 5
-proximal PArp ORF and a 3
-proximal RdRp ORF,
and expression of the RdRp as a fusion protein via

Fig. 2. (a) Phylogenetic tree of the L (RdRp) proteins of PeVA (shown in bold) and cytorhabdoviruses (CR) and
nucleorhabdoviruses (NR), with rabies virus (GenBank accession no. NC_001542) as the outgroup. (b) Phylogenetic tree
of the RdRp of PeLV (shown in bold), dsRNA viruses of three new genera (NG1–3) proposed by Spear
et al. (2010) and one megabirnavirus (MB), with Penicillium chrysogenum virus (GenBank accession no. AF296439) as the outgroup. Predicted amino acid sequences were aligned and analysed using clustal w (Thompson
et al., 1994) with 1000 bootstrap replicates. Bootstrap percentages above 70 % are shown at the nodes. Partial sequences are indicated by asterisks. Scale bar, 0.1 aa substitutions per site. BYSMV, barley yellow striate mosaic virus (GenBank accession no. FJ665628); NCMV, northern cereal mosaic virus (NC_002251); SCV (AY250986); LNYV (NC_007642); LMoV (NC_011532); MMV, maize mosaic virus (NC_005975); TaVCV, tare vein chlorosis virus (NC_006942); PYDV, potato yellow dwarf virus (NC_016138); RYSV, rice yellow stunt virus (NC_003746); SYNV, Sonchus yellow net virus (NC_001615); MFSV, maize fine streak virus (NC_005974); CuYAV (X92203); CiTV1 (GU979420); SpFV1 (GU979419); PgV2, Phlebiopsis gigantea virus 2 (AM111097); DsRV1, Diplodia scrobiculata RNA
virus 1 (EU547739); FgV3, Fusarium graminearum virus 3 (GQ140626); RnMBV1, Rosellinia necatrix megabirnavirus 1
(NC_013462); LeV, Lentinula edodes virus HKB (AB429554); PgV1, Phlebiopsis gigantea virus 1 (AM111096).

1920
translational frameshifting (Spear et al., 2010). All these criteria are based on the properties of CiTV1 and SpFV1, whose genomes have been completely determined (Spear et al., 2010). PeLV had a genomic organization similar to CiTV1, SpFV1 and some other fungal dsRNA viruses (Spear et al., 2010). Spear et al. (2010) also proposed three new genera, all of which had the suggested defining criteria and could be classified by the threshold for percentage nucleotide sequence identity. A phylogenetic tree of the RdRp of PeLV and dsRNA viruses in the three putative new genera (Spear et al., 2010) and one megabirnavirus showed that PeLV formed a distinct cluster with CiTV1, SpFV1 and CuYAV (Fig. 2b) and may thus have the same defining properties of these dsRNA viruses.

The 5’-proximal ORF of PeLV contained levels of proline and alanine (total 19.3%) as high as other PArps (Fig. 1d) (Spear et al., 2010). The FSFinder algorithm (Moon et al., 2004) implicated a shifty heptanucleotide motif (GGGAAAC at nt 3549–3555) and a pseudoknot around the PArp ORF stop codon of PeLV, which could be expressed downstream of the RdRp ORF via −1 ribosomal frameshifting as a PArp–RdRp fusion protein (Fig. 1d). The sequence of PeLV did not appear to encode a putative coat protein (Fig. 1d). Investigation of actual properties for lack of genome encapsidation, having a dsRNA genome and expression of RdRp as a fusion protein, and biological properties other than graft transmissibility should be performed. At the least, PeLV may be a virus species that is closely related to CiTV1, SpFV1 and CuYAV (Fig. 2b).

Acknowledgements

The authors thank Drs A. Sato and A. Kono for persimmons, Dr A. Azuma for nucleic acids extraction and Dr Nakaune for technical advice.

References


