Title: Exotic and indigenous viruses infect wild populations and captive collections of temperate terrestrial orchids (*Diuris* species) in Australia

Authors: Stephen J. Wylie, Hua Li, Kingsley W. Dixon, Helen Richards, Michael G.K. Jones

PII: S0168-1702(12)00373-5
DOI: doi:10.1016/j.virusres.2012.10.003
Reference: VIRUS 95842

To appear in: *Virus Research*

Received date: 6-7-2012
Revised date: 2-10-2012
Accepted date: 3-10-2012

Please cite this article as: Wylie, S.J., Li, H., Dixon, K.W., Richards, H., Jones, M.G.K., Exotic and indigenous viruses infect wild populations and captive collections of temperate terrestrial orchids (*Diuris* species) in Australia, *Virus Research* (2010), doi:10.1016/j.virusres.2012.10.003

Exotic and indigenous viruses infect wild populations and captive collections of temperate terrestrial orchids (*Diuris* species) in Australia

Stephen J. Wylie[1]*, Hua Li[1], Kingsley W. Dixon[2], Helen Richards[3], Michael G.K. Jones[1]

[1] Plant Virology Section, Plant Biotechnology Research Group, Western Australian State Agricultural Biotechnology Centre, School of Biological Sciences and Biotechnology, Murdoch University, Perth, Western Australia 6150, Australia.

[2] Botanic Gardens and Parks Authority, Kings Park and Botanic Garden, West Perth, Western Australia 6005, Australia.

[3] Australian Orchid Foundation, PO Box 322, Essendon North, Victoria 3041, Australia.

*corresponding author

Stephen J. Wylie

Plant Virology Section, Plant Biotechnology Research Group, Western Australian State Agricultural Biotechnology Centre, School of Biological Sciences and Biotechnology, Murdoch University, Perth, WA 6150, Australia.

Phone: +61 89360 6600

Fax: +61 89360 6303

Email: s.wylie@murdoch.edu.au

Abstract

Four species of *Diuris* temperate terrestrial orchids from wild and captive populations were tested for the presence of polyadenylated RNA viruses. The genomes of three exotic viruses were determined: two potyviruses, *Bean yellow mosaic virus* and *Ornithogalum mosaic virus*, and the polerovirus *Turnip yellows virus*. The genomes of five indigenous viruses were detected, including four novel species. They were the potyvirus Blue squill virus A, another potyvirus, two proposed capilloviruses, and a partitivirus. Partitivirus infection is of interest as this group of viruses is also associated with endophytic fungi (mycorrhizae) that are necessary for the germination, growth, development of many terrestrial orchids. Sequence divergence data indicate post-european, pre-european, and endemic origins for these viruses via inoculum from introduced and native plants. The implications of the findings of this study for orchid conservation, and particularly reintroduction programs where viruses may be spread inadvertently to wild populations from infected propagation sources, are discussed.

1. Introduction

The family *Orchidaceae* is one of the three largest groups of flowering plants in the world, with more than 25,000 species described from over 800 genera with orchid hybrids being major horticultural crops (Roberts and Dixon, 2008). While most tropical

and subtropical orchids are epiphytic, some are soil dwelling, and southern Australia is a centre of biodiversity for temperate terrestrial (geophytic) species (Jones, 1993; Brown et al., 2008).

Research in Australia revealed that a range of viruses, both indigenous and exotic, infect native orchids (Mackenzie et al., 1998; Gibbs et al., 2000). In the most comprehensive study, 850 individual Australian native orchid plants representing 72 genera were tested using generic primers (Gibbs et al., 2000). A mixture of exotic and indigenous viruses from five genera (*Potexvirus, Potyvirus, Rhabdovirus, Tobamovirus, Tospovirus*) was found. Some exotic viruses of commercial orchids (*Orchid fleck virus*, *Cymbidium mosaic virus*, and others) were found infecting native orchids, but other viruses (*Diuris virus Y*, *Ceratobium mosaic virus*, Pterostylis virus Y) were previously unknown and may represent endemic Australian groups (Mackenzie et al., 1998; Gibbs et al., 2000).

The terrestrial orchid genus *Diuris* was named after the long lateral sepals on the flowers, although the distinctive ear-like petals have given them their common name of Donkey Orchid. Over 50 species of *Diuris* orchids have been described, and all but one are endemic to Australia. While some *Diuris* species are relatively common, others are threatened, often as a result of habitat loss. Activities such as land clearing, browsing by introduced herbivores, competition by weeds, and infection by introduced pathogens all threaten some populations and species.

Here, we describe investigations to identify viruses infecting *Diuris* orchid plants growing in natural habitats and those propagated in *ex situ* collections by non-government organizations and conservation agencies for re-introduction programs into populations where plants are diminished in number or have become extinct. We describe the viruses found, consider whether they represent recent or long-standing associations with their hosts, and speculate on implications for orchid conservation. The *Diuris* species studied were *D. magnifica* D.L.Jones (pansy orchid) (Jones, 1991), *D. corymbosa* Lindl. (common donkey orchid) (Lindley, 1840), and *D. laxiflora* Lindl. (bee orchid) (Lindley, 1840), which are not considered threatened in the southwest of Western Australia where they are endemic (Hoffman and Brown 2011). The fourth species, *D. pendunculata* R.Br. (small snake orchid) (Brown, 1810) is endemic to two regions in the north-east and south-east of New South Wales, where it is listed as endangered under the Federal Environment Protection and Biodiversity Conservation Act 1999, the New South Wales Threatened Species Conservation Act 1995, and under Appendix II of CITES (Convention on International Trade in Endangered Species).

2. Materials and Methods

2.1. *Plant materials and RNA extraction*

Total RNA was extracted from samples of eighteen *Diuris* plants: five *D. magnifica* from a major *ex situ* conservation orchid collection in Western Australia, three *D.*

pendunculata plants from a private conservation collection from Victoria, Australia, and five each of *D corymbosa* and *D. laxifolia* plants from two remnant forest sites near the town of Brookton, Western Australia. Total RNA was extracted using an RNeasy Plant RNA kit (Qiagen). Total RNA was quantified and its quality measured (RNA integrity number >6.5) using a Nanodrop ND-1000 spectrophotometer (NanoDrop Technologies) and Bioanalyszer 2100 (Agilent) before 5 µg RNA from each plant was pooled.

2.2. *Sequencing and analysis*

Purification of polyadenylated RNA from total RNA using Oligo-dT labeled beads, nebulization of RNA prior to adaptor ligation, library construction, amplification, and single-end sequencing of 101 nt reads using Illumina HiSeq2000 technology was done by Macrogen Inc, Seoul, South Korea, and paired-end sequencing of 90 nt reads using the same technology by Beijing Genomics Institute (BGI), Shenzhen, China. Three Illumina sequencing reactions were done. Reaction KP (Macrogen Inc) was with two *D. magnifica* plants from the *ex situ* population in Perth, Western Australia. Reaction SW3.1 (BGI) was with one wild *D. laxiflora* and three wild *D. corymbosa* plants sampled from a remnant forest site near Brookton township, Western Australia. Sequencing reaction SW3.3 (BGI) was done with one wild *D. laxiflora* plant sampled from a remnant forest site near Brookton, and one *D. pendunculata* plant from an *ex situ* population in Victoria, Australia (Table 1).

106 *De novo* assembly of contigs was carried out using CLC Genomics Workbench v4.8
107 (www.clcbio.com) and Geneious Pro v5.5.6 (Drummond et al., 2012). Parameters for the
108 assembly of contigs were minimum overlap of 50% of read length (50 nt), 10%
109 maximum gaps per read, and independent assemblies were done using 85%, 90% and
110 95% minimum overlap identity. Resulting contigs were sorted according to length and
111 those less than 2 kb were removed. Batches of remaining contigs were subjected to
112 Blastn, Blastx and Blastp analysis against appropriate GenBank databases. Putative virus
113 sequences identified were edited manually where necessary to remove gaps and
114 determine where aberrant reads were sequencing errors, i.e. when aberrant bases occurred
115 1-5% of the time. Final virus consensus sequences were constructed using contigs from
116 both assemblers. Open reading frames and identities of deduced proteins, mature
117 peptides, and domains encoded by them were predicted within Geneious Pro, the NCBI
118 Conserved Domain Database (CDD), InterProScan
119 (http://www.ebi.ac.uk/Tools/pfa/iprscan), and by identity after alignment with
120 characterized virus sequences. Pairwise identities were calculated after alignment of
121 nucleotide (nt) and amino acid (aa) sequences using ClustalW in Geneious Pro.
122 Phylogenetic trees of aa sequences were constructed using Neighbor-joining (NJ),
123 Maximum Likelihood (ML), and Maximum Parsimony (MP) methods in MEGA5 [16]
124 after ClustalW pairwise alignment of sequences with a Gonnet protein weight matrix, a
125 gap open penalty of 10, a gap separation distance of 4, and a gap extension penalty of 0.2.
126 Tests of phylogeny were done using 1000 bootstrap replications. Host plants were

matched with viruses using RT-PCR assays with virus-specific primers, as described previously (Wylie et al., 2012a), followed by Sanger sequencing of amplicons. Primers (5'>3'; F = forward; R = reverse) used to detect viruses in plants were as follows: BSVAF, AGCGGCAGTAGGGGCTCAGT; BSVAR, AACATCTCTTGCCCAACCCACCA; BYMVF, GCGCGTGAGCAACAGAAGCG; BYMVR, CCTCCCCCAAGCACTCCTGC; DiVAF, CCCCAACCTGTGGGGGCTCT; DiVAR, TTCCGAACAAACGGGGCGGG; DiVBF, ACAGTCGGATTCCCGGGCGA; DiVBR, AGCTTCCGCGGACAGGGACT; DOVAF, CGAACCGACACGACCGCTGG; DOVAR, TTGCACCAGTGCCCTGAGCC; DPCVF, CGCTGTCGCCCGGTTCACTT; DPCVR, GACGCC CCAGATGGTTCGCA; OrMVF, TTTCGAGGCCTTCGTGCGGC; OrMVR, TTCACGCGACCA ACTCGCCC; TuYVF, AGAGTTTTACTATCCCGCAAAGC; TuYVR, TGCTTCTGCGTAGAG ATTCTCG. Total RNA was extracted from plants as described above. First strand cDNA synthesis was done using the reverse primer in the presence of Improm-II™ reverse transcriptase (Promega). PCRs were carried out over 40 cycles in the presence of GoTaq® polymerase (Promega) using an annealing temperature of 60^{o}C. Amplicons were sequenced with the primers used to create them using the Sanger method (AB 3730) after purification through MinElute® columns (Qiagen).

3. Results

Three sequence data sets were obtained after Illumina sequencing. Two datasets of

148 26,753,454 and 26,468,430 reads were of 90 nt paired-end reads, and one dataset of
149 18,668,558 was of 101 nt single-end reads. *De novo* assembly of each dataset was done
150 independently using two software packages and contig size was limited to >2,000 nt.
151 Contig sets for each assembly were assembled against one another to identify duplicate
152 sequences that could be removed. After this analysis, 18,661 contigs remained. When
153 these were analysed through Blastn and Blastx against appropriate databases, the majority
154 (18,499) had >60% identity with sequences of eukaryotic origin, usually plants, or
155 transposable elements. The remainder matched prokaryotic sequences, had no matches
156 on GenBank (orphans), or matched plant viral sequences. Orphan sequences were
157 analysed for conserved virus-like motifs such as RNA-dependent RNA polymerases
158 (RdRp), but none was identified.
159
160 Eleven contigs had virus-like motifs or sequence identity with described plant viruses.
161 After manual editing to remove sequencing errors at the ends of reads where necessary,
162 the genomic architecture of each sequence was predicted by identifying possible open
163 reading frames (ORF), analysis against the NCBI Conserved Domain Database (CDD),
164 and InterProScan. Eleven viral sequences belonging to distinct isolates of eight species
165 belonging to four families were identified from the eight *Diuris* plants tested. Using
166 specific primers in RT-PCR assays followed by sequencing of amplicons, one virus
167 isolate was identified in each of seven plants tested, and four virus isolates were
168 identified in one plant (Table 1).

169

170 3.1. *Potyvirus*

171 The genome sequences of seven isolates of four potyvirus species were detected from
172 three *Diuris* species in both *ex situ* and wild populations (Table 1).

173

174 3.1.1 *Blue squill virus A*

175 A sequence representing the complete genome of Blue squill virus A (BSVA) isolate
176 SW3.1 was from a wild plant of *D. corymbosa* (Table 1). The consensus sequence was
177 assembled from 155,838 reads (0.582% total reads) with a coverage range of 66-6,263X
178 across the genome and mean coverage of 1,425X. It was subsequently detected in another
179 plant in the same population using an RT-PCR-based assay with specific primers. Both
180 plants had mild leaf mottling visible on young leaves. There were no flowers or immature
181 flower spikes on infected plants. Of the uninfected *D. corymbosa* plants growing nearby,
182 only approximately one third of them had flowers or flower spikes. The BSVA sequence
183 shared 87% nt and 95% aa identities with the near-complete genome sequence of BSVA-
184 KP1 described previously from the Australian non-orchidaceous monocot plant
185 *Chamaescilla corymbosa* (Blue Squill) (*Asparagaceae*) (Wylie et al., 2012a). These high
186 sequence identities confirm that the new isolate is a member of the BSVA species
187 (Adams et al., 2005). BSVA is closely aligned with other potyviruses within the *Bean*
188 *common mosaic virus* (BCMV) subgroup that are identified only from Australia,
189 *Hardenbergia mosaic virus*, *Passionfruit woodiness virus* (Fig. 1) (Gibbs et al., 2008;

Wylie and Jones, 2011a, b; Wylie et al., 2012a). The new isolate shared 71-77% aa identity with them across their complete polyprotein sequences (Table 2). Genome organization of BSVA was typical of potyviruses. There was one large ORF encoding a polyprotein of 3064 aa and calculated mass of 349.6 kDa. The ORF begins at AUG (nt 260-262) and terminates at UAA (nt 9521-9523). The polyprotein is processed into 10 mature proteins, and an ORF within the P3 domain +2 frame encoded the PIPO (Table 3) (Chung et al., 2008). This represents the first complete genome sequence of an isolate of BSVA.

3.1.2. *Bean yellow mosaic virus*

Two similar sequences were detected in wild *D. corymbosa* and *ex situ D. magnifica* plants that exhibited quite severe leaf and flower spike distortion and patches of dead cells on the leaves (Fig. 2; Table 1). Blastn and Blastx analyses showed that the sequences closely resembled the genomes of *Bean yellow mosaic virus* (BYMV) isolates from international locations and from a range of host species, including orchids (Wylie et al., 2008). The genomes of the two new BYMV isolates shared 97% aa identity and 95% nt identity with one another (Table 2). The sequence of the genome of BYMV isolate SW3.1 was constructed from 2,523 reads (0.009% total reads) with a coverage range of 3-49X and mean coverage of 24X. The sequence of BYMV isolate KP2 was constructed from 11,121 reads (0.059% total reads), with a coverage range of 12-388X and mean coverage was 118X. When the complete polyprotein sequences were aligned with those

211 of the six other BYMV isolates for which complete genomes were available, the new
212 isolates shared closest aa identity (94-97%) with isolate Fr (GenBank accession
213 FJ492961) from *Freesia* plants collected in South Korea. They shared least aa identity
214 (76-77%) with isolate CS (AB373203) from *Pisum sativum* in Japan. The coat protein
215 (CP) sequences shared 98-99% identity with isolates of BYMV from an Australian
216 legume *Kennedia prostrata* (GenBank accessions DQ901434, DQ901435), the
217 agricultural legume *Lupinus angustifolius* (narrow-leafed lupin) originally from the
218 Mediterranean but cultivated in Australia (GenBank accessions AF192781, AF192782,
219 EU082121), and *Gladiolus* plants from India (GenBank accession AM398198) and Japan
220 (GenBank accession AB041972). Genomic organization was typical of that of BYMV
221 isolates and potyviruses generally (Table 3). The wild plant of *D. corymbosa* in which
222 BYMV isolate SW3.1 was detected grew in a patch of remnant native vegetation on a
223 road verge surrounded by clover pastures and lupin crops, both of which are reservoirs of
224 BYMV (Wylie et al., 2008). The captive *D. magnifica* plant where BYMV isolate KP2
225 was detected showed characteristic leaf mottling and patches of necrosis. When species-
226 specific primers were used to screen five other *D. magnifica* plants in the population that
227 exhibited similar symptoms, and five that did not show symptoms, the virus was present
228 in all of the symptomatic plants and three of the asymptomatic plants.
229
230 3.1.3. *Ornithogalum mosaic virus*

Three sequences shared 98.0% nt and 98.2% aa identity with one another. They were identified as genomes of isolates of OrMV (Table 1) by their sequence identity (81-82%) with a near-complete genome sequence of OrMV from *Iris* in Australia (JN127345) (Fig. 1, Table 2). The isolates were detected in wild *D. corymbosa* and *D. laxifolia* plants and in a captive *D. magnifica* plant, all of which had symptoms typical of virus infection (Table 1). The number of reads that corresponded to OrMV isolates KP1, SW3.1, and SW3.3 were 9,002 (0.048% total reads), 23,256 (0.086%) and 6,333 (0.023%) respectively. The coverage ranges were 7-433X, 18-1,144X, and 7-321X, respectively. Mean coverage was 96X, 222X, and 60X respectively. A screening of other captive plants of the same species using virus-specific primers showed that the virus was present in 5/5 symptomatic and 0/5 asymptomatic plants tested. Blastp analysis of the CP sequences revealed they shared high aa identity (98-99%) with partial genome sequences of isolates of Pterostylis virus Y (PtVY), a virus first described from *Pterostylis* and *Eriochilus* orchids endemic to southeastern Australia (Gibbs et al., 2000). The sequences shared high % identity to *Ornithogalum mosaic virus* (OrMV) isolates from around the world, including those from orchids such as *Vanilla*. When all 32 available complete OrMV and PtVY coat protein (CP) sequences were aligned, highest aa identities (96-97%) were with OrMV isolates from *Lachnalia* in South Africa (GenBank accession FJ159371) and USA (GenBank accession FJ159373), and from *Ornithogalum* in Japan (GenBank accession AB079647). Least aa identities (65%) were with CPs of isolates from *Gladiolus* in India (e.g. GenBank accession JN692498), the low identity indicating

252 these "OrMV" sequences probably belong to another species. Genome organization of
253 the three new OrMV isolates was almost identical to one another and typical of potyvirus
254 genomes. The deduced polyprotein was 3016 aa with a calculated mass of 341.5 kDa.
255 The ORF began at an AUG start codon (nt 125-127) and terminated at UGA (nt 9167-
256 9169). The PIPO was in the +2 frame at nt 2747-2967 within the P3 and encoded a
257 putative protein of 73 aa with a mass of 8.5 kDa (Table 3). These represent the first
258 complete genome sequences of isolates of OrMV.
259
260 3.1.4. *Donkey orchid virus A*
261 A sequence of 9,867 nt representing the complete or near-complete genome of a novel
262 potyvirus was identified from a wild plant of *D. laxiflora* that did not exhibit symptoms
263 (Table 1). The genome sequence was determined by assembly of 2,286 reads (0.008% of
264 total reads) with a range of coverage of 2-65X and with mean coverage of 21X. The
265 deduced polyprotein sequence was approximately equidistant from other potyviruses for
266 which a complete polyprotein sequence was available, it shared 38-41% aa identity with
267 them (Fig. 1, Table 2). The name Donkey orchid virus A (DOVA), isolate SW3.1 was
268 applied. Its predicted genome organization was typical of those of other potyviruses,
269 consisting of 10 gene products within the polyprotein, and a small ORF encoding the
270 PIPO protein located in the +2 frame within the P3. The polyprotein encoded a
271 polyprotein of 3183 aa with a calculated mass of 360 kDa. The ORF began at an AUG
272 start codon (nt 118-120) and terminated at UGA (nt 9664-9666). Mature protein positions

and sizes within the polyprotein were typical of potyviruses (Table 3). The putative PIPO ORF encoded a protein of 72 aa with a calculated mass of 8.4 kDa. The PIPO gene began at the conserved GGAAAAAA motif (nt 3186-3193) (Chung et al., 2008) and terminated at a UGA stop codon (nt 3404-3406). There was no AUG translation initiation codon present in the PIPO gene. The aa sequence of the DOVA CP shared highest identity, 56-61%, with isolates of BYMV (e.g. AB041971), tobacco vein banding virus (e.g. GU904046), and potato virus Y (e.g. AY742733). Although the viral vector was not identified, the highly conserved aphid transmission motifs of RITC (nt 1204-1214) and PTK (nt 2290-2298) at the N- and C-terminal regions, respectively, of the HC-Pro, and the DAG motif (nt 8890-8898) at the N-terminus of the CP were present, indicating that it is most likely transferred from plant to plant via aphids. Unlike most other potyviruses thought to have evolved in Australia (Gibbs et al., 2008; Webster et al., 2008), DOVA is not closely aligned with potyviruses of the BCMV subgroup (Fig. 1).

3.2. *Betaflexivirus*

Analysis of two virus-like sequences from two captive-grown *D. pendunculata* plants determined they shared 65% nt identity. Blast analyses suggested they were genomes of betaflexiviruses. Nucleotide identity between the sequences was below the species demarcation point (72%) for flexiviruses (Adams et al., 2004), so they were given the species names Diuris virus A (DiVA) and Diuris virus B (DiVB). The genome sequences of DiVA and DiVB were assembled from 125,394 (0.473% total reads) and 25,133 reads

294 (0.094%), respectively. Coverage range was 29-6,852X and 12-1,385X and mean
295 coverage was 1,626X and 323X, respectively. Each of the genomes shared a similar
296 degree of nt identity, 62-63%, with that of Hardenbergia virus A (HarVA, GenBank
297 accession NC_015395). HarVA is an unassigned member of the family *Betaflexiviridae*
298 recently identified from an Australian endemic legume, *Hardenbergia comptoniana*
299 (Wylie and Jones 2011c). Although HarVA is not currently assigned a genus, its genome
300 organization and phylogeny aligns it closely with *Cherry virus A* (CVA), a member of the
301 genus *Capillovirus* (Jelkman 1995), and more distantly to *Apple stem grooving virus*
302 (ASGV), the type species of the genus (Wang et al., 2010). The genomic organizations of
303 DiVA and DiVB were predicted through comparison with the genomes of HarVA, CVA
304 and ASGV, and by analysis with InterProScan and CDD. DiVA and DiVB shared highly
305 similar genome architectures with one another and with HarVA (Fig. 3). In common
306 with HarVA and the two capilloviruses, the DiVA and DiVB genomes were distinctive in
307 that they lacked an in-frame stop codon at the C-terminus of the replicase. The extent of
308 the replicases of both DiVA and DiVB genomes were estimated to be from nt 39-5003,
309 encoding a protein of 1655 aa with a calculated mass of 192.5 kDa. Like HarVA and
310 capilloviruses, in DiVA and DiVB replicases there were four domains: methyltransferase
311 (Met); papain-like protease (P-Pro); UvrD helicase (Hel); and an RNA-dependent RNA
312 polymerase (RdRp). Conserved motifs were present in the same positions for each of the
313 new viruses, and they were highly similar to those of HarVA (Wylie and Jones, 2011c).
314 As with HarVA, the conserved Met motif DEAD/H was present as DECH (nt 561-572)

315 (Rozanov et al., 1992; Hirata et al., 2010). The calalytic Cys residue (Koonin and Dolja,
316 1993) in the P-Pro domain was present at nt 2073-2075. Conserved Hel motifs A (GKS)
317 and B (DE) (Koonin and Dolja, 1993) were identified at nt positions 2631-2639 and
318 2826–2831, respectively. The conserved RdRp core motif S/TGx3 Tx3 NS/Tx22 GDD
319 (Koonin, 1991) was present as TGx3 Tx3 NTx22 GDD at nt 4548-4655. Replicase
320 sequences of DiVA and DiVB shared 66% nt and 61% aa identities, confirming they are
321 distinct species. Complete replicase aa sequences were aligned with homologous
322 sequences of isolates of alpha-, beta-, and gamma- flexiviruses, and NJ, MP and ML
323 analyses were done. All phylogenetic methods placed DiVA, DiVB, and HarVA closely
324 together, and the capillovirus CVA clustered nearby (Fig. 4). Closest aa identities with
325 the replicase sequences of DiVA and DiVB were with those of CVA (28%), ASGV
326 (26%), Cherry mottle leaf virus (*Trichovirus*) (26% DiVA, 28% DiVB), Apple chlorotic
327 leaf virus (*Trichovirus*) (26% DiVA, 27% DiVB) and Banana mild mosaic virus
328 (unassigned a genus) (26% DiVA, 28% DiVB), and the lowest of 14% identity was with
329 the fungal alpha- and gammaflexiviruses *Sclerotinia sclerotiorum debilitation-associated*
330 *RNA virus* **(***Sclerodarnavirus*) and *Botrytis virus F* (*Mycoflexivirus*), respectively**.** The
331 extent of the CPs of DiVA and DiVB were predicted by comparison with CPs of
332 betaflexiviruses, and by analysis on CDD. The start codons for CPs were predicted to
333 begin at the methionine located at nt 6018-6020 (DiVA) and nt 6021-6023 (DiVB) in the
334 contexts AGCCAUGGU and GAACAUGG, respectively. Translation from there yields
335 a protein of 232 aa with a calculated mass of 26 kDa for each virus, which closely

matches the masses of CPs of ASGV strains (26.3- 27.1 kDa). CP sequences of DiVA and DiVB shared 70% nt and 76% aa identities. CPs of HarVA shared 68% nt and 74% aa identity with DiVA and 71% nt and 78% aa identity with DiVB. These figures are close to the flexivirus species demarcation point for CPs (<72% nt and <80% aa identites for CPs or replicases) proposed by Adams et al. (2004), and confirm that DiVA, DiVB, and HarVA are distinct, although closely related virus species.

A 30K-like movement protein (MP) occurred in the +2 reading frame where its C-terminus overlapped the N-terminal region of the predicted CP (Fig. 4). In both DiVA and DiVB the protein was initiated with AUG at nt 5045-5047. In DiVA the extent of the MP was 389 aa with a predicted mass of 43.8 kDa, terminating at UGA. In DiVB it was slightly smaller at 380 aa with a predicted mass of 42.7 kDa, terminating at UAA (Fig. 4). The conserved MP motif DxR was present at nt 5381-5389 as DGR in both viruses. The MPs of DiVA and DiVB shared 64% nt and 53% aa and identity with one another. The MP of DiVA shared 56% identity with that of HarVA, and DiVB shared 48% identity with it. They shared low identity with capilloviruses CVA (DiVA 23%, DiVB 18%) and ASGV (DiVA 16%, DiVB 22%). Type members of the betaflexivirus genera *Citrivirus*, *Tepovirus*, *Trichovirus* and *Vitivirus*, that also have 30K-like MPs, shared 17-24% aa identity with the MPs of DiVA and DiVB.

3.3. *Partitivirus*

Two sequences of 2,010 and 1,806 nt were detected from the plant of *D. pendunculata* tested, the same plant co-infected with DiVA and DiVB. The consensus sequence of the larger sequence was assembled from 3,346 reads (0.012% total reads) with a range of coverage of 12-398X and mean coverage of 150X. The smaller sequence was assembled from 2,805 reads (0.010% total reads) with a range of coverage of 15-305X and mean coverage of 140X. Blastx analysis revealed that the large segment shared highest aa identity (64%) with the complete replicase segment of an unclassified member of the *Partitiviridae*, *Heterobasidion RNA virus 5* (HetRV5) (GenBank accession HQ541326), and lower identity with other partitiviruses (Fig. 5). The nt sequence of the smaller segment shared very little identity with any available sequence, but the aa sequence of the single putative ORF shared 30% identity with the CP of *Heterobasidion RNA virus 1* (Genbank accession HQ541324), and 24% identity with the CP of *Amaysa cherry disease-associated mycovirus* (GenBank accession NC_006440). Partitiviruses are bisegmented double-stranded RNA viruses, with one segment encoding a coat protein and the other a replicase. The large segment of the new virus encodes the putative replicase and the smaller the CP. The replicase sequence had one predicted ORF of 1866 nt that started at AUG (nt 83-85) and ended at UGA (nt 1946-1948). The ORF encodes a RdRp-like protein of 622 aa with a calculated mass of 73 kDa. The RdRp core motif S/TGx3 Tx3 NS/Tx22 GDD (Koonin, 1991) was present as SGx3 Tx3 DSx26 GDD (nt 1283-1402). This RdRp motif was conserved amongst the 14 partitiviruses analyzed with the exceptions of *Sclerotinia sclerotiorum partitivirus S* and *Raphanus sativus cryptic*

virus 1, where there were 27 and 28 aa residues between the DS and GDD motifs, respectively. A polyadenylation signal as such was not present. Instead, the 3' end of the genome segment has an AU-rich sequence (nt 1965-2010) where adenosines accounted for 61% of the nucleotides and uracils 24%. This AU-rich sequence, called the 'interrupted' poly (A) tail, is reported in other partitiviruses (e.g. Lim et al., 2005). The putative CP segment has one predicted ORFflanked on either side by a 5' UTR of 103 nt and a 3'UTR of 212 nt. The 3' UTR of the CP segment has an interupted poly (A) tail, in this case of 60 nt (nt 1747-1806). The ORF of 1,491 nt begins at AUG (nt 104-106) and ends at UGA (nt 1592-1594). It encodes a protein of 497 aa with MW 54.7 kDa. Hence the sequence identities and predicted gene functions of these two segments of RNA indicate that they represent the complete genome of a partitivirus, designated as Diuris pendunculata cryptic virus (DPCV) (Table 1).

3.4. *Polerovirus*

A sequence of 5629 nt was isolated from the *D. pendunculata* plant co-infected with DiVA, DiVB and DPCV. The whole nt sequence shared 89% identity with the complete genome sequence of the type isolate of *Turnip yellows virus* (TuYV-FL1), and 80% nt identity with complete genome sequences of *Brassica yellows virus* (BrYV). The consensus sequence was assembled from 2,031 reads (0.007% total reads) with a coverage range of 4-224X and mean coverage of 150X. The sequence is approximately the same size as the genome of TuYV-FL1 (5641 nt). As with isolate TuYV-FL, the new

399 sequence is predicted to have six ORFs (ORF0-ORF5), a 5'UTR of 31 nt, an intergenic
400 non-coding region of 203 nt (nt 3209-3411) between ORF2 and ORF3, and a 3' UTR of
401 85 nt. The 5'-terminal nts were ACAAAA, identical to those of many poleroviruses. One
402 of the species demarcation criteria for viruses of the family *Luteoviridae* is that
403 differences in aa sequences of any gene product should be greater than 10% (D'Arcy and
404 Domier, 2005). Five deduced protein sequences encoded by the new TuYV genome
405 shared greater than 90% aa identity with homologous sequences of TuYV-FL1, but the
406 read-through protein (ORF5) shared slightly less, at 87% identity. With the exception of
407 the read-through domain (RTD), all proteins in this isolate met the species identity
408 criterion, so it is proposed that this new sequence represents the genome of an isolate of
409 TuYV. The isolate was designated as TuYV-SW3.1 (Table 1).
410

411 **4. Discussion**

412 Eleven complete or partial genome sequences of eight virus species from four *Diuris*
413 species, both wild and captive plants, were identified. Of the virus species discovered,
414 three - BYMV, OrMV, TuYV - are clearly recent arrivals into Australia because they
415 have international distribution in other hosts. The other five species represent endemic
416 taxa that may have evolved in Australia. BSVA was recently described from a non-
417 orchidaceous indigenous plant (Wylie et al., 2012a). Four others had not been described,
418 although they concur with established families with members distributed internationally.
419 Some infected plants exhibited symptoms typical of virus infection, whilst others

420 remained symptomless. Notably, a plant of the rare species *D. pendunculata* was co-
421 infected with four viruses from three families, including one exotic virus, yet appeared
422 asymptomatic.
423
424 4.1. *Exotic and indigenous potyviruses*
425 Potyviruses are an extremely successful genus, having global distribution from a wide
426 phylogenetic spread of host plants. Many aphid species are vectors. It is suggested that
427 potyviruses adapted to agricultural expansion over 7000 years ago in the fertile crescent
428 and subsequently spread and evolved as humans carried plant propagules with them to
429 new lands (Gibbs and Ohshima, 2010). Several potyviruses are described only from
430 Australia, including HarMV and BSVA, and phylogeny places most in the BCMV
431 subgroup. Nucleotide identities of these "Australian BCMV-like potyviruses" is usually
432 within the range of 75-77% (Wylie and Jones, 2011a), figures that hover very close to the
433 species demarcation point for potyviruses (<75-76%) (Adams et al., 2005), indicating
434 they have a recent common ancestor that evolved as it encountered new hosts (Gibbs et
435 al., 2008). Recent plant-virus associations generally induce more severe symptoms than
436 older associations, which may be of a mutualistic nature (Roossinck, 2011). It is
437 estimated from nucleotide substitution rates in potyviruses that BCMV-like potyviruses
438 arrived in Australia only about 2000 years ago, possibly transported by Austronesians on
439 their crop plants as they colonized the Pacific (Gibbs et al., 2010). Indeed, both HarMV
440 and BSVA induce symptoms in natural host species, suggestive that these are quite recent

441 associations (Wylie and Jones, 2011a, Wylie et al., 2012a). The potyvirus DOVA is
442 unusual because its nucleotide sequence reveals that it does not belong to the BCMV
443 subgroup and it did not induce symptoms on the host plant. It is not closely allied to any
444 known potyvirus group, signifying that it has evolved in isolation. It's polyprotein
445 sequence is approximately equidistant (38-41% aa identity) from all other known
446 potyviruses, signifying that it has probably evolved in isolation for a long period. Along
447 with another Australian orchid virus in the *Potyviridae*, the proposed poacevirus
448 Caladenia virus A (Wylie et al 2012b), its presence is an indication that potyviruses may
449 have invaded Australia on more than one occasion during pre-history, and perhaps at a
450 much earlier time than previously speculated (Gibbs et al., 2010), thereby enabling more
451 mature (asymptomatic) relationships between host and virus to develop. Conversely, it
452 may represent a recent incursion by an unknown potyvirus. Obtaining sequence data
453 from further DOVA isolates should address this question because greatest sequence
454 diversity exists in the region where a species evolved (Thresh, 1980; Gibbs et al., 2008).
455
456 In our study, exotic viruses BYMV and OrMV induced severe symptoms typical of virus
457 infection, including necrotic patches on the leaves, whereas the other exotic virus, TuYV,
458 did not. The presumed indigenous viruses did not induce visible symptoms.
459
460 The presence of the exotic potyviruses BYMV and OrMV infecting both wild and captive
461 orchid populations is of concern, and further studies to determine host ranges and

462 incidences amongst other orchid species should be undertaken. BYMV has been isolated
463 from several orchid species internationally, including the Australian species *Pterostylis*
464 *curta* (GenBank accession AF185960) (Gibbs et al., 2000), and from a wide range of
465 non-orchidaceous species (Wylie et al., 2008). The new isolates closely resemble those
466 described from introduced leguminous crops in the region, and these are the likely origin
467 of the virus. Infected *Diuris* plants exhibited quite severe symptoms of infection (leaf
468 and flower stem distortion, chlorosis) that may reduce host fecundity and lifespan, but
469 this has not been proven experimentally. Less clear is the origin of OrMV isolates. The
470 virus was recently identified for the first time in Western Australia from cultivated
471 (introduced) *Iris* plants growing in a home garden (Wylie et al., 2012a). Other members
472 of the *Iridaceae* are also hosts (Wei et al., 2006), notably *Gladiolus* (e.g. GenBank
473 accession FJ573184). Several iridaceous genera including *Gladiolus, Homeria, Ixia*, and
474 *Freesia* are naturalized weeds of African origin widespread in Australian forests and
475 shrubland, but these have not been tested as hosts of OrMV. The presence of highly
476 identical OrMV sequences from widely distributed *Diuris* populations indicates that a
477 single introduction into the region has occurred. Isolates with high % identity from
478 *Pterostylis* and *Eriochilus* orchids in eastern Australia were given the name Pterostylis
479 virus Y (Gibbs et al., 2000), and this name should be changed to *Ornithogalum mosaic*
480 *virus*.
481
482 4.2. *Australian flexiviruses*

High sequence identity and similar genome organizations of DiVA, DiVB and HarVA place them together with members of the genus *Capillovirus*, especially to CVA. We propose that Diuris virus A, Diuris virus B, and Hardenbergia virus A are placed together with ASGV and CVA in the genus *Capillovirus*. The Australian capilloviruses are genetically equidistant from one another, suggesting simultaneous divergence from a common ancestor. Together with Scaevola virus A, a member of the closely related genus *Trichovirus* (Martelli et al., 2007; Hirata et al., 2010) described from a non-orchidaceous species of the Australian flora (Wylie et al., 2012a), these viruses hint at a long presence of flexiviruses on the Australian continent.

4.3. *Partitiviruses and orchids*

To our knowledge, the partitivirus DPCV is the first to be identified in Australia and the first from orchids. Partitiviruses are persistent dsRNA viruses with bisegmented genomes that are vertically transmitted through the gametes, but not *via* grafts, manual inoculations or vectors (Blanc, 2007). Partitiviruses infect fungi, plants and protozoa. Phylogenetic analysis of the partitivirus RdRp genes suggests that they have been transmitted between plants and fungi (Roossinck, 2010). This is of particular interest in orchids as they are dependent on fungal associations to germinate and grow. Many Australian terrestrial orchids have reduced root systems and so are highly dependent on mycorrhizal associations for mineral nutrition (Brundrett, 2007). In most cases, orchids associate with basidiomycetes (Smith and Read, 2008). The DPCV sequence clustered closest to that of

504 the basidiomycete fungus virus HetRV5. In another plant/fungus/dsRNA virus system,
505 all three partners were required in a mutualistic association that imparted heat tolerance to
506 a grass that grew in an active thermal area (Marquez et al., 2007). Partitivirus-like partial
507 CP and RdRp sequences are found stably integrated in some plant genomes (Chiba et al.,
508 2011), but we consider it unlikely that DPCV is an example of integration because its
509 sequence was transcribed and the complete replicase RNA segment was obtained.
510
511 4.4. *Exotic polerovirus*
512 TuYV is an exotic virus described from Tasmania, Australia, in cultivated brassicas and
513 peas, but until now has not been described from orchids. Luteovirids are restricted to the
514 phloem tissue of host plants and are transmitted by aphids in a persistent manner (Grey
515 and Gidow, 2003). This virus's presence is alarming because the host is a rare species and
516 any impact on plant growth by the virus may adversely affect the long-term survival of
517 the species.
518
519 4.5 *Method used*
520 Although detection of only polyadenylated RNA viruses was expected using this method
521 (oligo-dT bead purification of mRNAs prior to library construction), the partitivirus
522 RdRp sequence was determined despite its lack of a poly (A) tail *sensu stricto*. It is
523 probable that non-polyadenylated RNA viruses, viroids, and DNA viruses also co-infect
524 the tested plants, but these could not be detected with the method used. Others have

525 enriched samples for viral sequences before library construction by extracting dsRNA
526 (e.g. Roossinck et al., 2010), partially purifying virus particles (e.g. Thapa et al., 2012),
527 and sequencing short-interfering RNAs (siRNAs) (e.g. Kreuze et al., 2009). Addition of
528 genetic tags (barcodes) to sample sequences before pooling (e.g. Roossinck et al., 2010)
529 would simplify the process of matching viruses with host plants, although in this case
530 where there were a small number of samples sequenced per sequencing reaction, the
531 method used of designing specific primers to match virus with host in RT-PCR and
532 Sanger sequencing assays was simple and efficient.
533
534 Sequence coverage was generally lowest at the 5' and 3' ends of the genomes. Peaks
535 of coverage occurred in places along the genomes, seemingly independent of the
536 genome sequence. For example in the three ORMV isolates sequenced no pattern of
537 coverage was observed, suggesting that the patterns of peaks and troughs of
538 coverage observed is an artefact of library preparation or sequencing..
539
540 4.6 *Implications for conservation*
541 This study of only eight *Diuris* orchid plants of four species revealed a surprising number
542 and variety of RNA viruses, indicative that a rich indigenous viral flora, as well as an
543 aggressive exotic viral flora, exists in Australia's temperate terrestrial orchids, and in its
544 flora generally.
545

546 The presence of exotic viruses is of particular concern. BYMV and OrMV were present
547 in wild and captive populations, so the source of infection to the captive population was
548 probably *via* infected wild tubers. This study highlights the need to quarantine newly
549 collected propagules until plants are examined for symptoms, and ideally assayed for
550 viruses before introducing them to *ex situ* conservation or horticultural collections. In
551 captive populations where plants are grown in pots on benches, conditions exist for rapid
552 spread of viruses by vectors. The case of TuYV is of particular concern because the virus
553 is probably exotic and visible symptoms of infection were not apparent.
554
555 Elimination of exotic viruses from threatened orchid populations is certainly desirous for
556 their long-term survival. Except in the probable case of the partitivirus, it is uncertain
557 whether any of the identified viruses are transmitted through seed (Wong et al., 1994),
558 and this should be examined, as seed is a possible route to generate virus-free plants. Heat
559 and chemical treatments and *in-vitro* meristem culture have also been used to establish
560 virus-free orchid stocks (Lim et al., 1993).
561
562 It is not known what impacts the newly found indigenous viruses have on plant growth
563 and development. Indeed, it is possible some represent beneficial associations (Marquez
564 et al., 2007; Roossinck, 2011), and as such should be maintained. This study highlights
565 the need to gain a greater understanding of the roles viruses play in long-term
566 partnerships with plants, particularly for plants of conservation concern.

Importantly, conservation programs for orchids have considered virus infection, but usually only associated with epiphytic taxa. This research highlights the need to study the ecological roles of indigenous and exotic viruses as well as assess phyto-health risks to conservation of *ex situ* and wild orchid populations.

Acknowlegements

This study was funded by Australian Research Council Linkage Grant LP110200180 with support from the Western Australian Botanic Gardens and Parks Authority and the Australian Orchid Foundation.

References

Adams, M.J., Antoniw, J.F., Fauquet, C.M., 2005. Molecular criteria for genus and species discrimination within the family *Potyviridae*. Arch. Virol. 150, 459–479.

Adams, M.J., Antoniw, J.F., Bar-Joseph, M., Brunt, A.A., Candresse, T., Foster, G.D., Martelli, G.P., Milne, R.G., Fauquet, C.M., 2004. The new plant virus family *Flexiviridae* and assessment of molecular criteria for species demarcation. Arch. Virol. 149, 1045–1060.

Blanc, S,. 2007. Virus transmission–getting in and out, in: Waigman, E., Heinlein,

588 M.(Eds.),Viral transport in plants, Springer, Berlin, pp. 1–28.

589

590 Brown, R., 1810. *Diuris pedunculata.* Prodromus Florae Novae Hollandia et Insulae Van

591 Diemen, 316.

592

593 Brown, A., Dixon, K., Hopper, S., Dundas, P., 2008. Orchids of Western Australia,

594 University of Western Australia Press, Perth.

595

596 Brundrett, M.C., 2007. Scientific approaches to Australian temperate terrestrial orchid

597 conservation. Austral. J. Bot. 55, 293-307.

598

599 Chiba, S., Kondo, H., Tani, A., Saisho, D., Sakamoto, W., Kanematsu, S., Suzuki, N.,

600 2011. Widespread endogenization of genome sequences of non-retroviral RNA viruses

601 into plant genomes. PLoS Pathog 7: e1002146. doi:10.1371/journal.ppat.1002146.

602

603 Chung, B. Y-W., Miller, W.A., Atkins, J.F., Firth, A.E., 2008. An overlapping essential

604 gene in the *Potyviridae*. PNAS 105, 5897–5902.

605

606 Crawford, L.J., Osman, T.A.M., Booy, F.P., Coutts, R.H.A., Brasier, C.M., Buck, K.W.,

607 2006. Molecular characterization of a partitivirus from *Ophiostoma himal-ulmi*. Virus

608 Genes 33, 33-39.

609

610 D'Arcy, C.J., Domier, L.L., 2005. *Luteoviridae*, in: Fauquet, C.M., Mayo, M.A.,

611 Maniloff, J., Desselberger, U., Ball, L.A., (Eds.) Virus Taxonomy: VIIIth Report of the

612 ICTV, Elsevier Academic Press, London, pp. 891–900.

613

614 Drummond, A.J., Ashton, B., Cheung, M., Heled, J., Kearse, M., Moir, R., Stones-Havas,

615 S., Thierer, T., Wilson, A., 2012 Geneious v. 5.6.4, http://www.geneious.com.

616

617 Gibbs, A., Mackenzie, A., Blanchfield, A., Cross, P., Wilson, C., Kitajima, E.,

618 Nightingale, M., Clements, M., 2000. Viruses of orchids in Australia: their identification,

619 biology and control. Austral. Orchid Rev. 65, 10-21.

620

621 Gibbs, A.J., MacKenzie, A.M., Wei, K.-J., Gibbs, M.J., 2008. The potyviruses of

622 Australia. Arch. Virol. 153, 1411-1420.

623

624 Gibbs, A.J., Fargette, D., Garcia-Arenal, F., Gibbs, M.J., 2010. Time – the emerging

625 dimension of plant virus studies. J. Gen. Virol. 91, 13-22.

626

627 Gibbs. A.J., Ohshima, K., 2010. Potyviruses and the digital revolution. Ann. Rev.

628 Phytopathol. 48, 205-223.

629

630 Gray, S., Gildow, F.E., 2003. Luteovirus-aphid interactions. Ann. Rev. Phytopathol. 41,
631 539-566.
632
633 Hirata, H., Yamaji, Y., Komatsu, K., Kagiwada, S., Oshima, K., Okano, Y., Takahashi,
634 S., Ugaki, M., Namba, S., 2010. Pseudo-polyprotein translated from the full-length ORF1
635 of a capillovirus is important for pathogenicity, but a truncated ORF1 protein without
636 variable and CP regions is sufficient for replication. Virus Res. 152, 1–9.
637
638 Jelkmann, W., 1995. Cherry virus A: cDNA cloning of dsRNA, nucleotide sequence
639 analysis and serology reveal a new plant capillovirus in sweet cherry. J. Gen. Virol. 76,
640 2015-2024.
641
642 Jones, D.L., 1991. *Diuris magnifica.* Austral. Orchid Res. 2, 58.
643
644 Jones, D.L., 1993. Native Orchids of Australia. Imago Publishing Ltd., Hong Kong.
645
646 Koonin, E.V., Dolja, V.V., 1993. Evolution and taxonomy of positive-strand RNA
647 viruses: implications of comparative analysis of amino acid sequences. Crit. Rev.
648 Biochem. Mol. Biol. 28, 375–430.
649

650 Kozlowska-Makulska, A., Guilley, H., Szyndel, M.S., Beuve, M., Lemaire, O., Herrbach,
651 E., Bouzoubaa, S., 2010. P0 proteins of European beet-infecting poleroviruses display
652 variable RNA silencing suppression activity. J.Gen. Virol. 91, 1082–1091.
653
654 Kreuze, J.F., Perez, A., Untiveros, M., Quispe, D., Fuentes, S., Barker, I., Simon, R.,
655 2009. Complete viral genome sequence and discovery of novel viruses by deep
656 sequencing of small RNAs: a generic method for diagnosis, discovery and sequencing of
657 viruses. Virol. 388, 1–7.
658
659 Lim, S.T., Wong, S.M., Goh, C.J., 1993. Elimination of cymbidium mosaic virus and
660 odontoglossum ringspot virus from orchids by meristem culture and thin section culture
661 with chemotherapy. Ann. Appl. Biol. 122, 289-297.
662
663 Lim. W.-S., Jeong, J.H., Jeong, R.-D., Yoo, Y.B., Yie, S.W., Kim, K.-H., 2005. Complete
664 nucleotide sequence and genome organization of a dsRNA partitivirus infecting
665 *Pleurotus ostreatus.* Virus Res. 108, 111-119.
666
667 Lindley, J., 1840. *Diuris corymbosa*. Sketches of Vegetation of the Swan River Colony,
668 51.
669

Mackenzie, A.M., Nolan, M., Wei, K.-J., Clements, M.A., Gowanlock, D., Wallace, B.J., Gibbs, A.J., 1998. Ceratobium mosaic potyvirus: another virus from orchids. Arch. Virol. 143, 903-914.

Márquez, L.M., Redman, R.S., Rodriguez, R.J., Roossinck, M.J., 2007. A virus in a fungus in a plant – three way symbiosis required for thermal tolerance. Science 315, 513–515.

Martelli, G.P., Adams, M.J., Kreuze, J.F., Dolja, V.V., 2007. Family *Flexiviridae*: a case study in virion and genome plasticity. Ann. Rev. Phytopathol. 45, 73-100.

Roberts, D. L. & Dixon, K. W. 2008. Orchids. Curr. Biol. 18, R325-R329.

Roossinck, M.J., 2010. Lifestyles of viruses. Phil. Trans. Royal Soc. B 365, 1899-1905.

Roossinck, M,J., Saha. P., Wiley, G.B., Quan, J., White, J.D., Lai, H., Chavarr´ia, F., Shen, G., Roe, B.A., 2010. Ecogenomics: using massively parallel pyrosequencing to understand virus ecology. Mol. Ecol. 19 (Suppl. 1), 81–88.

Roossinck, M.J., 2011. The good viruses: viral mutualistic symbioses. Nature Rev. Microbiol. 9, 99-108.

Rozanov, M.N., Koonin, E.V., Gorbalenya, A.E., 1992. Conservation of the putative

693 methyl-transferase domain: a hallmark of the 'Sindbis-like' supergroup of positive-strand
694 RNA viruses. J. Gen. Virol. 73, 2129–2134.
695
696 Smith, F.F., Brierly, P., 1944. Ornithogalum mosaic. Phytopathol. 34, 497-503.
697
698 Smith, S.E., Read, D.J., 2008 Mycorrhizal symbiosis. Third Ed. Academic Press,
699 Elsevier, London.
700
701 Thapa, V., Melcher, U., Wiley, G.B., Doust, A., Palmer, M.W., Roewe, K., Roe, B.A.,
702 Shen, G., Roossinck, M.J., Wang, Y.M., Kamath, N., 2012. Detection of members of the
703 *Secoviridae* in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA. Virus Res.
704 167, 34-42.
705
706 Thresh, J.M., 1980. The origins and epidemiolology of some important plant virus
707 diseases. Appl. Biol. 5, 1–65.
708
709 Wei, T., Pearson, M. N., Cohen, D., 2006. First report of Ornithogalum mosaic virus and
710 Ornithogalum virus 2 in New Zealand. Plant Pathol. 55, 820.
711
712 Wang, L.P., Hong, N., Wang, G.P., Xu, W.X., Michelutti, R., Wang, A.M., 2010.
713 Distribution of apple stem grooving virus and apple chlorotic leaf spot virus in infected in

vitro pear shoots. Crop Prot. 29, 1447-1451.

Wong, S.M., Chong, C.G., Lee, Y.H., Tan, K., Zettler, F,W., 1994. Incidence of cymbidium mosaic and odontoglossum ringspot viruses and their significance in orchid cultivation in Singapore. Crop Prot. 13, 235-239.

Wylie, S.J., Coutts, B.A., Jones, M.G.K., Jones, R.A.C., 2008. Phylogenetic analysis of bean yellow mosaic virus isolates from four continents: relationship between the seven distinct groups found and their natural isolation hosts and geographical origins. Plant Dis. 92, 1596-1603.

Wylie, S.J., Jones, M.G.K., 2011[a]. Characterisation and quantitation of mutant and wild-type genomes of hardenbergia mosaic virus isolates co-infecting a wild plant of *Hardenbergia comptoniana*. Arch. Virol. 156, 1287-1290.

Wylie, S.J., Jones, M.G.K., 2011[b]. The complete genome sequence of passionfruit woodiness virus determined using deep sequencing, and its relationship to other potyviruses. Arch. Virol. 156, 479-482.

733 Wylie, S.J., Jones, M.G.K., 2011[c]. Hardenbergia virus A, a novel member of the
734 *Betaflexiviridae* from a wild legume in South-west Australia. Arch. Virol. 156, 1245-
735 1250.
736
737 Wylie, S.J., Luo, H., Li, H., Jones, M.G.K., 2012[a]. Multiple polyadenylated RNA viruses
738 detected in pooled cultivated and wild plant samples. Arch. Virol. 157, 271-284.
739
740 Wylie, S.J., Tan, A., Li, H., Dixon K.W., Jones M.G.K., 2012[b]. Caladenia virus A, an
741 unusual new member of the *Potyviridae* from terrestrial orchids in Western Australia.
742 Arch. Virol. (available online first)

743

744 **Figure Legends**

745

746 **Figure 1** Neighbor-joining tree of amino acid sequences of complete polyproteins of
747 potyviruses found in *Diuris* orchids (boxed) with other complete or near-complete
748 potyvirus polyproteins. GenBank accessions, virus name and isolate (if given) are shown.
749 Branch robustness was assessed by 1000 bootstraps shown as percentages. Those with
750 confidence values below 60% are not shown. Relationships calculated using Maximum
751 Parsimony and Maximum Likelihood methods gave essentially the same tree. The tree is
752 drawn to scale, with evolutionary distance used to infer the branch length in nucleotide
753 substitutions per site.

754

755 **Figure 2** Potted plants of *Diuris magnifica*. The plant on the right is infected with Bean
756 yellow mosaic virus isolate KP2. Visible is leaf distortion on the whole plant, and
757 chlorotic leaf mottle and necrotic streaks on the leaf (inset).

758
759

760 **Figure 3** Predicted genome organization of proposed species (A) Diuris virus A (DiVA)
761 and Diuris virus B (nucleotide positions that differ from DiVA in parentheses) compared
762 to (B) the genome organization of Hardenbergia virus A. Predicted nucleotide positions
763 of the borders of the regions described are given above or below the blocks representing
764 domains or genes, and the lengths of the untranslated regions (UTR) are given. Replicase

765 and coat protein (CP) genes of open reading frame 1 (ORF1) are represented by the upper
766 block. Shaded regions within ORF1 represent the methyltransferase (Met), papain-like
767 protease (P-pro), helicase (Hel), and RNA-dependent RNA polymerase (RdRp) domains
768 of the replicase gene (its predicted extent indicated by arrow). The movement protein
769 (MP) gene of ORF2 is represented in the lower block.
770
771 **Figure 4** Neighbor-joining tree of amino acid sequences of complete replicase proteins of
772 type isolates of 26 species within the families *Alpha-*, *Beta-*, and *Gammaflexivirus*. The
773 positions of the proposed species Diuris virus A and Diuris virus B (boxed) are shown in
774 relation to other species. GenBank accessions, species and genus names are shown.
775 Branch robustness was assessed by 1000 bootstraps shown as a percentage. Those with
776 confidence values below 60% are not shown. Relationships calculated using Maximum
777 Parsimony and Maximum Likelihood methods gave essentially the same tree. The tree is
778 drawn to scale, with evolutionary distance used to infer the branch length in nucleotide
779 substitutions per site.
780
781 **Figure 5** Neighbor-joining tree of amino acid sequences of complete replicase proteins of
782 14 partitiviruses isolated from plants and fungi. The proposed Diuris pendunculata
783 cryptic virus is boxed. GenBank accessions, species, and genus names are shown if
784 given. Branch robustness was assessed by 1000 bootstraps shown as a percentage. Those
785 with confidence values below 60% are not shown. Relationships calculated using

786 Maximum Parsimony and Maximum Likelihood methods gave essentially the same tree.

787 The tree is drawn to scale, with evolutionary distance used to infer the branch length in

788 nucleotide substitutions per site.

Table 1 Virus species and isolates identified in *Diuris* orchids.

Virus[a]	Isolate	Sequence length (nt)[b]	Classification Family, Genus	*Diuris* host species (plant/pot number)	Symptoms[c]	Host situation	Host location[d]	GPS location[e]	GenBank Accession number
BYMV	KP2	9533	*Potyviridae*, *Potyvirus*	*D. magnifica* (pot K43)	CLM, LD, FSD, NPL	Captive	Perth, WA	-31.955976, 115.842469	JX173278
BYMV	SW3.1	9530	*Potyviridae*, *Potyvirus*	*D. corymbosa* (plant B11)	CLM	Wild	Brookton, WA	-32.397438, 116.880024	JX156423
DOVA	SW3.1	9867	*Potyviridae*, *Potyvirus*	*D. laxiflora* (plant B9)	NVS	Wild	Brookton, WA	-32.397175, 116.880653	JX156422
BSVA	SW3.1	9842	*Potyviridae*, *Potyvirus*	*D. corymbosa* (plant B8)	MM	Wild	Brookton, WA	-32.397148, 116.880618	JQ807999
OrMV	KP1	9445	*Potyviridae*, *Potyvirus*	*D. magnifica* (pot K44)	CLM, FSD	Captive	Perth, WA	-31.955976, 115.842469	JQ807997
OrMV	SW3.1	9445	*Potyviridae*, *Potyvirus*	*D. corymbosa* (plant B10)	CLM	Wild	Brookton, WA	-32.397481, 116.879478	JQ807995
OrMV	SW3.3	9447	*Potyviridae*, *Potyvirus*	*D. laxifolia* (plant B22)	CLM	Wild	Brookton, WA	-32.396921, 116.881236	JQ807996
DiVA	SW3.3	6941	*Betaflexiviridae*, *Capillovirus*	*D. pendunculata* (pot 2081)	NVS	Captive	Yarra Glen, Vic	unknown	JX173276
DiVB	SW3.3	7001	*Betaflexiviridae*, *Capillovirus*	*D. pendunculata* (pot 2081)	NVS	Captive	Yarra Glen, Vic	unknown	JX173277
DPCV (Rep)	SW3.3	2010	*Partitiviridae*, *unclassified*,	*D. pendunculata* (pot 2081)	NVS	Captive	Yarra Glen, Vic	unknown	JX156423
DPCV (CP)	SW3.3	1806	*Partitiviridae*, *unclassified*,	*D. pendunculata* (pot 2081)	NVS	Captive	Yarra Glen, Vic	unknown	TBA
TuVY	WA-1	5629	*Luteoviridae*, *Polerovirus*	*D. pendunculata* (pot 2081)	NVS	Captive	Yarra Glen, Vic	unknown	JQ862472

a BYMV, bean yellow mosaic virus; DOVA, donkey orchid virus A; BSVA, blue squill virus A; OrMV, ornithogalum mosaic virus; DiVA, Diuris virus A; DiVB, Diuris virus B, DPCV (Rep), Diuris pendunculata cryptic virus, replicase segment; DPCV (CP), Diuris pendunculata cryptic virus, coat protein segment; TuVY, turnip yellows virus

b. Complete genome sequence length excluding poly (A) tail if present

c. CLM, chlorotic leaf mottle; FSD, flower spike distortion; LD, leaf distortion; MM, mild mottle; NPL, necrotic patches on leaves; NVS, no visible symptoms

d. Location of plants used in this study. WA, Western Australia; Vic, Victoria, Australia

e. Original locations of samples provided from an *ex situ* collection in Victoria could not be determined

Table 2 Pairwise identities (%) between amino acid sequences of complete and near-complete polyproteins of the potyviruses donkey orchid virus A, ornithogalum mosaic virus, blue squill virus A, and bean yellow mosaic virus found infecting *Diuris* orchids, and those of other potyviruses

Virus[a]	DOVA	OrMV SW3.1	OrMV SW3.3A	OrMV KP1	OrMV Bate9[b]	BSVA SW3.1	BSVA KP1[b]	BYMV SW3.1	HarMV	PWV	BCMV	SMV	WVMV	WMV	HiMV	PSbMV	SrMV	PenMV	PLMV	TNSV
OrMV SW3.1	41.2																			
OrMV SW3.3A	41.1	98.4																		
OrMV KP1	41.3	98.1	98.0																	
OrMV Bate9	40.1	81.9	82.3	81.6																
BSVA SW3.1	40.3	44.9	44.8	45.8	46.9															
BSVA KP1	40.7	45.1	45.1	44.9	46.9	94.5														
BYMV SW3.1	40.8	45.1	45.1	45.0	47.7	44.2	44.0													
HarMV	39.8	45.4	45.3	45.3	47.4	76.9	76.9	43.3												
PWV	40.2	44.4	44.5	44.6	46.6	71.2	71.1	44.3	71.3											
BCMV	39.5	45.2	45.2	45.3	47.1	64.2	66.4	45.3	64.3	65										
SMV	41.1	45.8	45.7	45.5	47.5	67.7	68.4	44.5	68.1	68.8	67.2									
WVMV	40.5	45.1	45.1	45.5	47.2	68.2	68.6	44.7	67.7	68.4	67.1	80.4								
WMV	39.1	45.1	44.9	45.3	47.1	65.3	67.7	44.2	64.9	65.2	68.6	81.8	76.6							
HiMV	37.8	42.9	43.4	43.5	45.8	40.7	41.2	41.3	41.5	41.4	39.7	41.5	41.6	39.8						
PSbMV	39.1	44.3	43.9	44.0	47.1	41.3	42.7	43.9	41.9	41.7	41.1	42.4	42.3	41	41.7					
SrMV	39.7	42.3	42.4	42.0	44.6	42.1	43.1	42.2	41.8	41.9	41.9	43.2	42.9	42.6	39.5	42				
PenMV	40.4	43.8	43.7	43.5	45.9	42.4	43.2	42.2	42.4	42.7	43.1	44	44.2	43.5	39.9	43.3	71.8			
PLMV	40.7	46.2	46.1	46.8	49.1	49.2	49.5	44.8	49.5	49.2	48.2	49.7	49.5	47.6	40.6	43.2	43	43.2		
TNSV	40.6	44.6	44.4	44.5	47.4	44	43.9	45.3	43.4	44	42.8	44.9	45.3	43	41.1	43.5	41.7	43.1	43.6	
PepMoV	39.8	44.3	44.3	44.2	46.5	43.9	43.9	44.8	43	43.5	42	43.3	44.1	41.8	41.2	42.5	41.5	41.7	43.7	57.7

[a] DOVA SW3.1 JX156422, Donkey orchid virus A isolate SW3.1; OrMV SW3.1, JQ807995 ornithogalum mosaic virus isolate SW3.1; OrMV SW3.3A, JQ807996 ornithogalum mosaic virus isolate 3.3A; OrMV-KP1, JQ807997 ornithogalum mosaic virus isolate KP1; OrMV Bate9, JN127345 Ornithogalum mosaic virus isolate Bate9; BSVA SW3.1, JQ807999 Blue Squill virus isolate SW3.1; BSVA KP1, JN052072 Blue Squill virus isolate SW3.1HarMV, HQ161081 Hardenbergia mosaic virus isolate 57.2; PWV, HQ122652 Passion fruit woodiness virus isolate MU-2; BCMV, AJ312437 Bean common mosaic

virus cowpea isolate R; SMV, FJ640972 Soybean mosaic virus isolate WS160; WVMV, AY656816 Wisteria vein mosaic virus isolate Beijing; WMV, HQ384216 Watermelon mosaic virus isolate Dendrobium; PSbMV, X89997 Pea seed-borne mosaic virus; HiMV, JQ395040 Hippeastrum mosaic virus isolate Marijiniup 1; TNSV, JQ314463 Tomato necrotic stunt virus isolate MX9354; PLMV, DQ851494 Peace lily mosaic virus isolate Haiphong; SrMV, AJ310198 Sorghum mosaic virus; PepMoV, EU586131 Pepper mottle virus isolate 205137; PenMV, DQ977725 Pennisetum mosaic virus isolate C

b near-complete polyprotein sequence

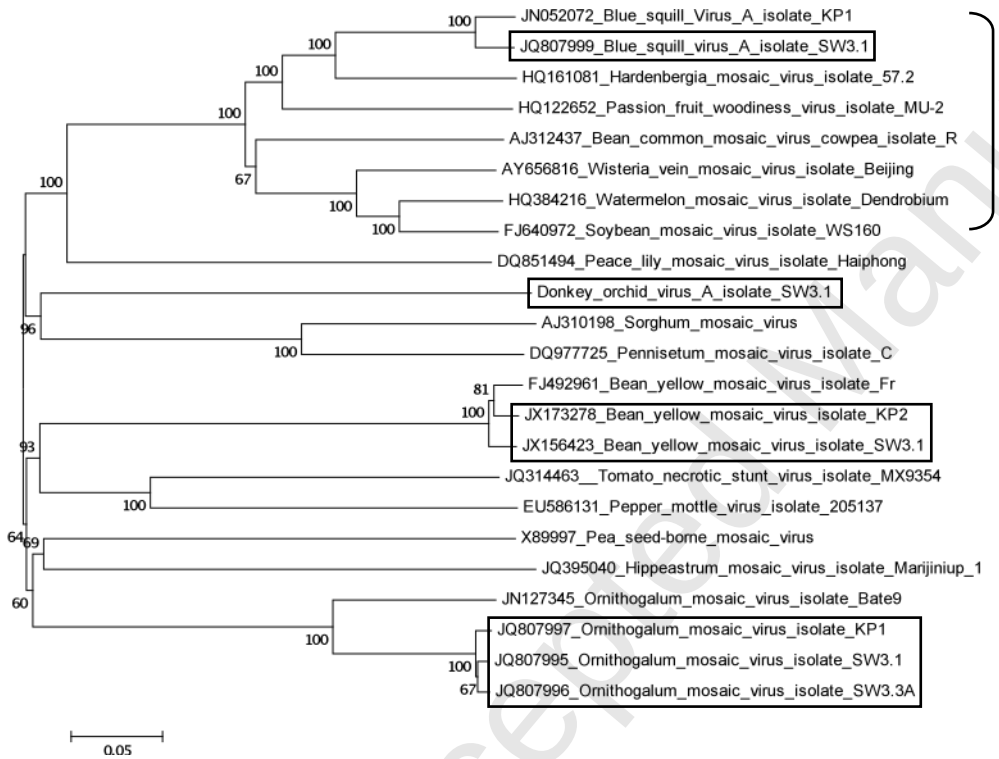
Table 3 Genome organization, predicted protein sizes and masses of new isolates of donkey orchid virus A, blue squill virus A, ornithogalum mosaic virus and bean yellow mosaic virus from *Diuris* orchids

Virus[a]	DOVA			BSVA			OrMV-KP1, SW3.1[b]			OrMV-SW3.3			BYMV-KP1			BYMV-SW3.1		
Genome feature[b]	nt	size aa	mass kDa	nt	size aa	mass kDa	nt	size aa	mass kDa	nt	size aa	mass kDa	nt	size aa	mass kDa	nt	size aa	mass kDa
Polyprotein	118-9666	3182	360.5	260-9523	3088	352.0	127-9174	3016	341.5	125-9169	3016	341.5	192-9362	3056	347.3	189-9440	3056	347.3
5'UTR	1-117	-	-	1-259	-	-	1-126	-	-	1-124	-	-	1-191	-	-	1-188	-	-
P1	118-1050	311	34.7	260-1210	317	36.4	127-927	267	29.6	125-925	267	29.6	192-1043	284	32.7	189-1040	284	32.7
HC-Pro	1051-2736	562	64.0	1211-2581	457	51.9	928-2292	455	51.3	926-2290	455	51.3	1044-2414	457	51.9	1041-2411	457	51.9
P3	2737-3777	347	40.2	2582-3628	349	40.3	2293-3336	348	40.5	2291-3331	347	40.2	2415-3458	348	40.3	2412-3455	348	40.5
PIPO	3186-3404	72	8.4	3034-3255	74	8.8	2752-2972	73	8.4	2747-2967	73	8.5	2876-3118	80	9.3	2873-3115	80	9.4
6K1	3778-3933	52	5.8	3629-3784	52	5.8	3337-3487	50	5.6	3332-3487	52	5.8	3459-3617	53	5.9	3456-3614	53	5.9
CI	3934-5850	639	71.6	3785-5686	634	71.4	3488-5394	635	70.9	3488-5389	634	70.7	3618-5522	635	71.2	3615-5519	635	71.2
6K2	5851-6009	53	5.8	5687-5845	53	6.1	5395-5553	50	5.6	5390-5548	53	5.9	5523-5681	53	6.0	5520-5678	53	6.0
NIa-VPg	6010-6588	193	21.8	5846-6415	190	21.8	5554-6129	192	21.5	5549-6124	192	21.5	5682-6254	191	22.1	5679-6251	191	22.0
NIa-Pro	6589-7314	242	27.4	6416-7144	243	27.8	6130-6858	243	27.5	6125-6853	243	27.4	6255-6983	243	27.2	6252-6980	243	27.1
NIb	7315-8871	519	59.3	7145-8695	517	59.5	6859-8412	518	59.5	6854-8407	518	59.4	6984-8540	519	59.0	6981-8537	519	59.0
CP	8872-9666	265	29.6	8696-9523	276	30.7	8413-9174	254	28.9	8408-9169	254	28.9	8541-9362	273	30.7	8538-9356	273	30.8
3'UTR	9667-9867	-	-	9524-9842	-	-	9175-9445	-	-	9170-9447	-	-	9363-9533	-	-	9357-9530	-	-

a DOVA, donkey orchid virus A; BSVA, blue squill virus A; OrMV, ornithogalum mosaic virus; BYMV, bean yellow mosaic virus

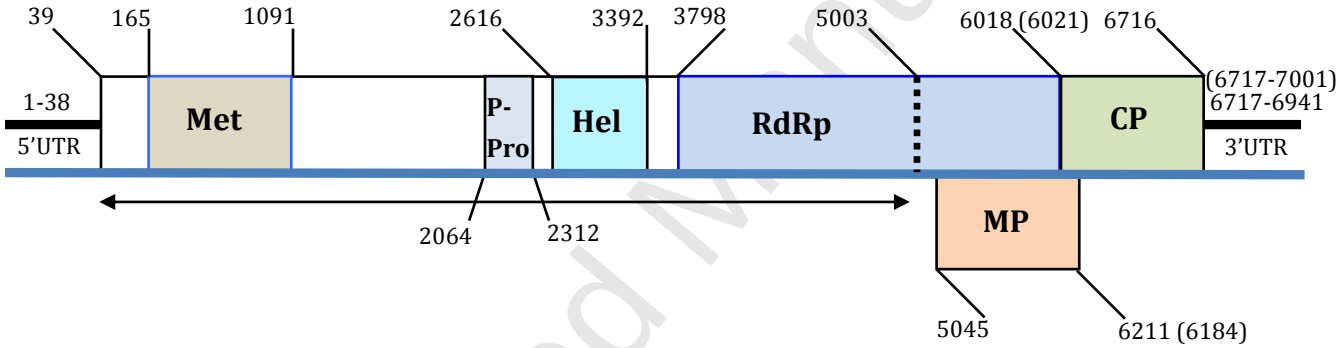
b Isolates OrMV-KP1 and OrMV-SW3.3 had identical genome organization so are presented together

b UTR, untranslated region; P1, protein 1; HC-Pro, helper component protease; P3, protein 3; PIPO, pretty interesting potyvirus ORF; 6K1, six kilodalton peptide 1; CI, cylindrical inclusion; 6K2, six kilodalton peptide 2; NIa-VPg, nuclear inclusion A, viral protein genome-linked; NIa-Pro, nuclear inclusion A protease; NIb, nuclear inclusion B; CP, coat (capsid) protein

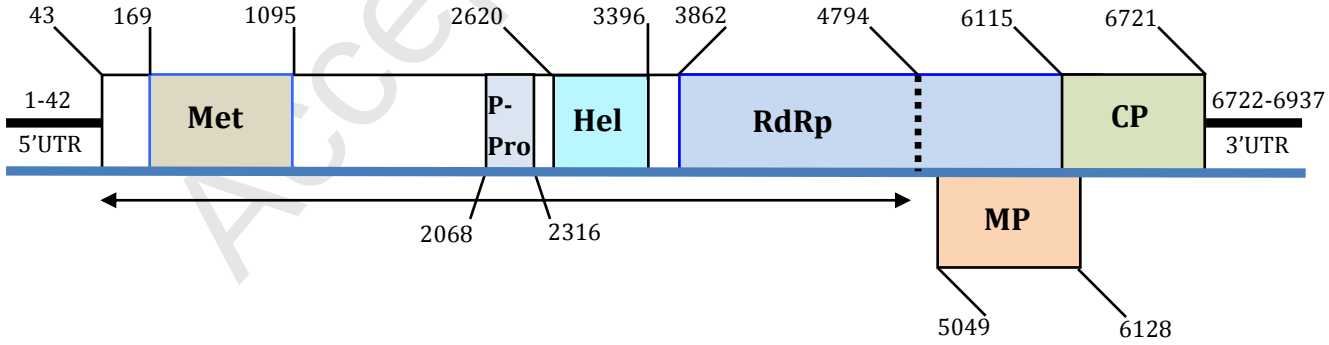


Figure(s)

A



B



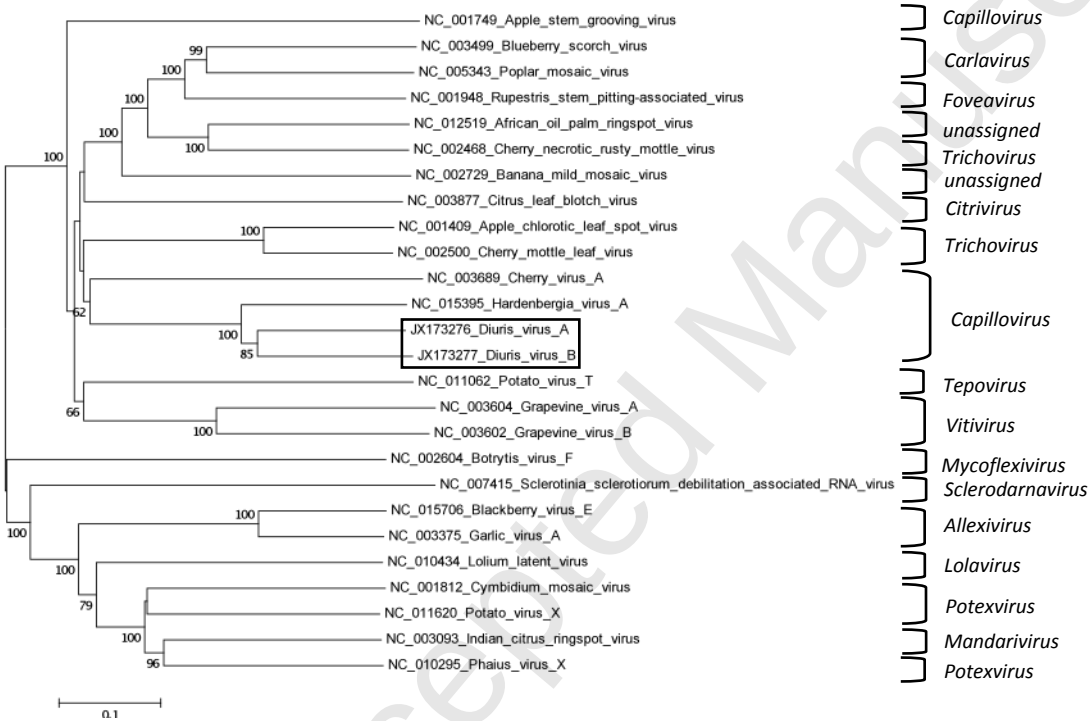


Fig 5