Accepted Manuscript

Title: Novel and divergent viruses associated with Australian orchid-fungus symbioses

Authors: Jamie W.L. Ong, Hua Li, Krishnapillai Sivasithamparam, Kingsley W. Dixon, Michael G.K. Jones, Stephen J. Wylie

PII: S0168-1702(17)30663-9
DOI: https://doi.org/10.1016/j.virusres.2017.11.026
Reference: VIRUS 97301

To appear in: Virus Research

Received date: 7-9-2017
Revised date: 21-11-2017
Accepted date: 24-11-2017

Please cite this article as: Ong, Jamie W.L., Li, Hua, Sivasithamparam, Krishnapillai, Dixon, Kingsley W., Jones, Michael G.K., Wylie, Stephen J., Novel and divergent viruses associated with Australian orchid-fungus symbioses. Virus Research https://doi.org/10.1016/j.virusres.2017.11.026

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Title:
Novel and divergent viruses associated with Australian orchid-fungus symbioses

Authors and affiliations:
Jamie W.L. Ong\textsuperscript{a}, Hua Li\textsuperscript{a}, Krishnapillai Sivasithamparam\textsuperscript{a}, Kingsley W. Dixon\textsuperscript{b},
Michael G.K. Jones\textsuperscript{a}, Stephen J. Wylie\textsuperscript{*}\textsuperscript{a}
\textsuperscript{a}Plant Biotechnology Group – Plant Virology, Western Australian State Agricultural Biotechnology Centre, School of Veterinary and Life Sciences, Murdoch University, Perth, Western Australia 6150, Australia
\textsuperscript{b}Department of Environment and Agriculture, School of Science, Curtin University, Bentley, Western Australia 6102, Australia

Corresponding author details:
Dr. Stephen J. Wylie
Address: Murdoch University, 90 South St, Murdoch, Western Australia 6150
Email: s.wylie@murdoch.edu.au; Tel: 61 8 9360 6600
Highlights

- Identified 22 novel viruses from wild terrestrial orchids and their fungal symbionts
- First reported virome of plants and their mycorrhizal partners in natural environment
- Two leaf-derived viruses were more related to mycoviruses than to plant viruses
- Extends known geographical range of Betapartitivirus, Hypovirus, Megabirnavirus, Mitovirus and Totivirus
- Inconsistent viral component associated with the same population over consecutive years

Abstract

Terrestrial orchids represent a symbiotic union between plants and mycorrhizal fungi. This study describes the occurrence and nature of viruses associated with one population of wild Pterostylis sanguinea orchids and their fungal symbionts over two consecutive years. A generic sequencing approach, which combined dsRNA-enrichment from plant and mycelial tissues, random amplification and high throughput sequencing, was used to identify novel viruses. The majority of the virus-like sequences represent partial genomes, and their identification is based solely on de novo assembly of sequencing data. In orchid leaf tissues we found three isolates of a novel totivirus and an unclassified virus; both resemble fungus-infecting viruses. Two isolates of Ceratobasidium sp were isolated from orchid underground stems contained at least 20 viruses, 16 of which were previously described as
alphapartitiviruses and betapartitiviruses. A hypovirus and a mitovirus were
genetically distant from existing members of the genera and did not readily fit into
recognised subgroups.

**Keywords:**
Mycorrhizal fungus; mycovirus; plant virus; *Pterostylis* orchid; virus taxonomy;
virome
1. Introduction

*Pterostylis* is a genus comprising of over 200 species terrestrial orchids indigenous to Australia, Indonesia, New Caledonia, New Zealand and Papua New Guinea (Janes & Duretto, 2010; Janes *et al.*, 2010; Brundrett, 2014). *Pterostylis* orchids have short roots ranging from 5-10 cm in length and they form obligate associations with compatible fungi that provide water and nutrients from beyond the rhizosphere (Ramsay *et al.*, 1986; Ramsay *et al.*, 1987). Australian terrestrial orchids differ from other composite organisms such as lichens in that the relationship is broken annually when the plant partner enters its dormant below-ground phase, and it becomes re-established when the shoot re-emerges, which may occur up to several years later, depending on environmental conditions (Brundrett, 2014). *Pterostylis* plants always establish mycorrhizal relationships with species of *Ceratobasidium* fungi (Warcup, 1973; Bonnardeaux *et al.*, 2007), but it is unclear if the same species or strain of fungus re-establishes the relationship each year.

The viruses of cultivated orchids are widely studied. The most common viruses are *Cymbidium mosaic virus* (CymMV) and *Odontoglossum ringspot virus* (ORSV) from families *Alphaflexiviridae* and *Virgaviridae*, respectively (Zettler *et al.*, 1990). Transmission vectors for these two viruses remain unknown but they can be spread through vegetative propagation and through trade of infected plants (Jensen, 1952; Blanchfield *et al.*, 2001). In contrast, viruses infecting wild orchids are far less well known. Exotic and indigenous viruses from five genera (*Dichorhavirus, Orthotospovirus, Potexvirus, Potyvirus* and *Tobamovirus*) were described from a mixture of captive and wild orchids in eastern Australia (Gibbs *et al.*, 2000). In Western Australia, wild orchids and their fungal symbionts were infected with exotic
and indigenous viruses of *Alphapartitivirus, Betapartitivirus, Divavirus, Endornavirus, Goravirus, Platypuvirus, Polerovirus* and *Potyvirus* (Wylie et al., 2012; Wylie *et al.*, 2013a; Wylie *et al.*, 2013b; Ong *et al.*, 2016a; Ong *et al.*, 2016b; Ong *et al.*, 2017). In Japan, wild *Calanthe izu-insularis* orchids were infected with cucumber mosaic virus (genus *Cucumovirus*) (Kawakami *et al.*, 2007). In India, isolates of ORSV (genus *Tobamovirus*), CymMV (genus *Potexvirus*) and a novel potyvirus infected wild epiphytic orchids (Sherpa *et al.*, 2006; Singh *et al.*, 2007).

In the current study, a generic approach based on high throughput sequencing was used to identify RNA viruses infecting plants of *Pterostylis sanguinea* (dark banded greenhood orchid) and mycorrhizal fungi associated with them. The samples were collected from a small natural population over two consecutive years. These orchids generate one or more new underground tubers each year, and the old tuber dies. Tubers typically germinate unevenly, with some remaining dormant from one to several years (Brundrett, 2014). Sixteen partitiviruses (of *Alphapartitivirus* and *Betapartitivirus*) were previously detected and characterised from the same orchid mycorrhizae (Ong *et al.*, 2017). In this report, we describe the novel non-partitiviral viruses identified from them and discuss them in ecological and evolutionary contexts.

2. Materials and methods

Leaves and underground stems were collected from *P. Sanguinea* orchids from their natural habitat in Western Australia. The subsequent experiments were carried out as described in Ong *et al.* (2017). Briefly, nucleic acids were extracted from the leaves and associated mycorrhizal fungi using a cellulose-based method (Morris &
DNA libraries were made from randomly primed cDNA. Amplicons were barcoded and pooled before high-throughput sequencing on an Illumina Hi-Seq platform. De novo assembly of contigs was done using CLC Genomic Workbench v6.5.1 (Qiagen), and subsequent analysis was done using Geneious v7.0.6 (Biomatters) and MEGA v6.06 software.

3. Results

3.1. De novo assembly

Three datasets of 153,581,198 and 92,046,118 and 35,630,376 reads, each of 101 nt, were generated from three independent Illumina sequencing runs. The first dataset included 4,875,054 reads from *Ceratobasidium* sp. F-2012. Samples F-2012 and P-2012 were sequenced as part of the second dataset and generated 2,854,974 and 5,725,698 reads respectively. F-2012 was re-sequenced to improve the coverage of the viral sequences. The third sequencing run generated 409,030 reads from P-2012 and 604,012 reads from F-2013. The remaining unaccounted reads were from other barcoded samples that are unrelated to this project. De novo assembly of the datasets generated a total of 119,854 contigs (12,896; 72,206 and 34,752 from each respective dataset) ranging from 200 nt to 22,685 nt, with a N50 length of 403, 359 and 305 respectively. Of these, 90 contigs were identified as virus-like; 14 were derived from *P. sanguinea* plants (474-10,716 nt) and 76 were from mycorrhizal fungi (407-8227 nt).

3.2. Identity of fungi

ITS sequences of both fungal isolates shared 99.7% nt identity, indicative they were of the same taxon of *Ceratobasidium* (Genbank accessions KU239992 (F-2012))
and KU239993 (F-2013)). The mycorrhizal fungi isolates were not identified to the level of species. The sequences were closest to a Ceratobasidium isolate from Pterostylis orchids from eastern Australia (GQ405561; e-value: 0.0, 91% coverage and 99% nucleotide identity) and to a Ceratobasidium-anamorph, Rhizoctonia sp. (JQ859901; e-value: 0.0; 100% coverage and 97% nucleotide identity) isolated from a diseased strawberry root in Western Australia.

3.3. Viruses from orchid-associated mycorrhizal fungi

3.3.1. Ceratobasidium mitovirus 1 (CbMV1): a proposed new mitovirus

A virus-like contig of 2850 nt was identified from Ceratobasidium in 2012 (Table 1). 82,998 sequence reads were mapped to it with pairwise identity of 84.8%, and the mean coverage of the proposed virus genome was 3573.9-fold. There was a single ORF (nt 235-2688) whose encoded protein product has a predicted mass of 92 kDa. An RdRp-like domain was identified at aa 228-543 (nt 916-1863) (Fig 1A), indicative of a replicase function. Further support that the single ORF encoded a replicase was the existence of six core RdRp motifs between aa 297-508 (nt 1123-1758) (Hong et al., 1999). The deduced protein sequence shared closest pairwise identities with replicases of mitoviruses (family Narnaviridae, genus Mitovirus), which infect the mitochondria of fungi (Hillman & Esteban, 2012). Mitovirus genomes typically comprise a single non-encapsidated positive-strand RNA of 2.3–2.9 kb, which encodes a single protein of 80–104 kDa believed to function as a replicase (Hillman & Esteban, 2012).

The genome sequence was most closely related to the mycorrhizal fungus-infecting mitoviruses Rhizophagus clarus mitovirus 1 (ReMV1-RF1) from Japan
(27% aa, 45% nt), Rhizophagus sp. HR1 mitovirus (RMV-HR1) from Japan (27% aa, 45% nt) and Tuber excavatum mitovirus (TeMV) isolated from Germany (20% aa, 43% nt), which together formed groups distinct from currently proposed mitovirus clades I and II (Fig 2A) (Doherty et al., 2006; Hillman & Cai, 2013). The deduced protein sequence shared 18-25% aa identity with other mitoviruses, figures below the accepted species demarcation limit of <40% (Hillman & Esteban, 2012). Thus, we propose that the sequence represents the complete genome of a previously-undescribed member of genus Mitovirus that we designate Ceratobasidium mitovirus 1 isolate Murdoch-1 (CbMV1; GenBank accession KU291923), named after the virus host genus and the location of its discovery.

3.3.2. Ceratobasidium virus 1 (CbV1): a proposed new mycovirus

Three partial monopartite virus sequences of 8227 nt, 7161 nt and 4051 nt respectively were identified from the Ceratobasidium isolate collected in 2012 (Fig 1B; Table 1). These sequences were designated Ceratobasidium virus 1 (CbV1) isolates Murdoch-2, Murdoch-3 and Murdoch-4 (GenBank accessions KU291947, KU291948 and KU291949 respectively). The sequence of isolate Murdoch-2 was constructed from 9593 raw sequence reads, with mean pairwise identity of 82.2% and a 119.6-fold mean coverage per base across the genome. 6792 reads were mapped to isolate Murdoch-3 with mean pairwise identity of 82.0% and mean coverage of 101.2-fold per base. Isolate Murdoch-4 was generated from 2558 reads of mean pairwise identity of 80.5% and mean coverage of 71.6-fold per base. CbV1 isolates Murdoch-2 and Murdoch-3 had two non-overlapping ORFs – a complete ORF1 (177 kDa) and partial ORF2 (RdRp; >74kDa and >115 kDa) while the sequence of Murdoch-4, which represented about 45% of its genome, encoded an incomplete RdRp (>130
Neither a ‘slippery sequence’ nor pseudoknot, typical of ribosomal frameshift sites (Brierley et al., 1992), was detected upstream of proposed ORF2. Comparison of the isolates showed 43-80% nt identity between genomes, 91% aa identity (80% nt) between the ORF1s, and 53-95% aa identity (59-81% nt) between the respective RdRps.

Blastp analysis showed that about 10% of the translated product of ORF1s of isolates Murdoch-2 and Murdoch-3 matched the transposase of Desulfovibrio oxyclinae (Table 1). The encoded CbV1 RdRps shared highest identity with RdRp of Rosellinia necatrix mycovirus 1-W1032/S5 (also named Yado-nushi virus W1032a; Zhang et al., 2016) at a pairwise identity of 31-33% aa and 47-50% nt. CbV1 and Rosellinia necatrix mycovirus 1-W1032/S5 were grouped together and distantly related to other unclassified mycoviruses (Fig 2B).

3.3.3. Ceratobasidium virus 2 (CbV2): a proposed new mycovirus

A contig of 7089 nt representing Ceratobasidium virus 2 (CbV2) isolate Murdoch-5 (GenBank accession KU291938) was constructed from 8508 reads, with mean coverage of 127.3-fold per base across the genome and mean pairwise identity of 46.1%. The partial genome of CbV2 encoded two non-overlapping ORFs representing a hypothetical protein (nt 1-3340; >117 kDa) and an RdRp (nt 3604-7089; >31 kDa) (Fig 1C). There was no evidence of ribosomal frameshift ‘slippery sequence’ sites upstream of the junction of the two putative ORFs. In ORF2, an RdRp_4 domain (pfam02123) was identified at aa 597-866 (nt 5392-6201) (Fig 1C) and the core RdRp motifs V and IV of T/SGx3 Tx3 NS/Tx22 GDD (where x is any
residue) (Koonin, 1991) were represented at aa 758-800 (nt 5875-6003) as SGx3 Tx3 NTx29 GDD.

Blast and phylogenetic analyses showed that CbV2 grouped most closely with an unclassified mycovirus Rhizoctonia solani RNA virus HN008 (RsRV-HN008; Zhong et al., 2015) (Fig 2B, Table 1). The two viruses shared 45% nt identity between genomes, 16% aa (43% nt) identity between ORF1s and 31% aa (48% nt) identity between ORF2s.

3.3.4. Ceratobasidium hypovirus 1 (CbHV1): a proposed new hypovirus

A contig of 7143 nt was generated from 5051 sequence reads with mean coverage of 74.7-fold and pairwise identity of 85.5%. The partial genome sequence has a 5’ untranslated region (UTR) of 89 nt and two predicted ORFs (Fig 1D). ORF1 was 3636 nt in length and is predicted to encode a protein of 133 kDa. The 3’ part of the genome was not obtained, and so ORF2 is incomplete (nt 3888-7143; >121 kDa). Based on its length compared to related hypoviruses and the position of RdRp domain, the partial ORF2 sequence represents about 60% of its complete ORF. Elements resembling slippery sequences and pseudoknots (Brierley et al., 1992) upstream of ORF2 were absent. The RdRp domain was located at aa 582-677 (nt 5631-5918) (Fig 1D) and the core RdRp motifs V and IV were represented at aa residues 590-634 (nt 5655-5789) as TGx3 Tx3 NTx31 GDD.

The virus represented by this sequence was designated Ceratobasidium hypovirus 1 (CbHV1) isolate Murdoch-6 (GenBank accession KU291924). We propose CbHV1 as a new hypovirus; the first hypovirus to be detected from a
basidiomycete host. CbHV1 shared highest aa identity (15-16%) with two definitive members of Hypovirus (Cryphonectria hypovirus 1-2, CHV1-2; family Hypoviridae) that infect Cryphonectria parasitica, the Chestnut blight fungus (Shapira et al., 1991; Hillman et al., 1994) and two other proposed members infecting Fusarium species (Fusarium graminearum hypovirus 1 (FgHV1) from China (Wang et al., 2013)) and Sclerotinia species (Sclerotinia sclerotiorum hypovirus 2 (SsHV2) from New Zealand (Khalifa and Pearson, 2014)). Hypoviruses are proposed to be categorised into subgroups Alphahypovirus and Betahypovirus, distinguished by having either two or one ORF, respectively (Nuss & Hillman, 2012; Yaegashi et al., 2012). A third subgroup, Gammahypovirus, has been tentatively proposed by Khalifa and Pearson (2014) to categorise SsHV2. Having two ORFs, CbHV1 was expected to share greater sequence identity with alphahypoviruses than betahypoviruses, but phylogenetic analysis positions CbHV1 distantly from members of each group (Fig 2C).

Comparison of CbHV1 with the hypoviruses showed similar nt identity at 41-43% and low aa identity; 12-16% aa identity with alphahypoviruses and 11-13% aa identity with betahypoviruses. This is consistent with the species demarcation limit of less than 60% aa identity between CHV-1 and CHV-2 and 50% aa identity between CHV-3 and CHV-4 (Nuss & Hillman, 2012).

3.4. Virus-like sequences identified from leaf samples

Four distantly related virus-like sequences were identified from P. sanguinea leaf tissue samples, one in 2012 and three in 2013. Because these viruses resembled mycoviruses, but not known plant viruses, attempts to amplify fungus sequences from the leaf samples using primers ITS1 and ITS4 were carried out. However this test did
not detect fungi from the leaf. This evidence suggested that these two mycovirus-like viruses may indeed be plant viruses.

3.4.1. Pterostylis sanguinea virus A (PsVA), a mycovirus-like virus from leaf tissue

A partial monopartite virus genome sequence of 10,716 nt was identified from leaf tissue in 2012 (Fig 1E; Table 1). 56,636 101 nt reads were mapped to the sequence, with a pairwise identity of 87.4%, and a mean coverage of 1444.8-fold across its partial genome. This sequence was named Pterostylis sanguinea virus A (PsVA) isolate Murdoch-7 (GenBank accession KU291925).

The PsVA genome has two consecutive non-overlapping ORFs of 5739 nt and 3758 nt, respectively (Fig 1E). Ribosomal frameshift was not detected in this sequence. PsVA is predicted to have an unusually long 5' UTR of 1154 nt. The first putative translational start codon is positioned at nt 1155 corresponding to the start of the hypothetical CP, estimated to have a mass of 208 kDa (Fig 1E). A region encoding a Nudix hydrolase-like domain, responsible for hydrolysis of nucleoside diphosphate derivatives, was detected at aa residues 177-297 (nt 1683-2045) (Fig 1E).

In ORF2, an RdRp domain, which corresponds to similar domains in viruses belonging of Chrysovirus, Luteovirus, Rotavirus and Totivirus, was detected at aa residues 934-1209 (nt 9758-10,585) (Marchler-Bauer & Bryant, 2004; Marchler-Bauer et al., 2013). The highly conserved core RdRp motifs V and VI (S/TGx3 Tx3 NS/Tx22 GDD) (Koonin, 1991) were present at aa 1109-1151 (nt 10,283-10,411) as SGx3 Tx4 NTx28 GDD.
ORF1 shared highest identity with the CP-encoding ORF1 of the monopartite mycovirus Lentinula edodes spherical virus (LeSV; an unclassified virus), identified from Shiitake mushroom (Lentinula edodes) in South Korea (Won et al., 2013). The PsVA RdRp showed highest identity to RdRps of Lentinula edodes mycovirus isolates HKA and HKB (LeV; unclassified), also from Shiitake mushroom but from Japan (Ohta et al., 2008; Magae, 2012), the saprophytic fungus virus Phlebiopsis gigantea large virus-1 (PgLV-1; unclassified; Kozlakidis et al., 2009) and mycorrhizal fungus infecting Thelephora terrestris virus 1 (TtV1; unclassified; Petrzik et al., 2016) (Fig 2B; Table 1). These unclassified viruses shared similar genome organisation of unipartite RNA genomes encoding two ORFs and are phylogenetically grouped within a new proposed genus, Phlegivirus (Fig 2B; Petrzik et al., 2016). With the exception of PsVA, all Phlgiviruses are detected from basidiomycetous hosts. ClustalW comparison of PsVA, LeV, PgLV-1 and TtV1 isolates showed 41-43% nt identity across the genomes. Identity between the homologous proteins of PsVA, LeV, PgLV-1 and TtV1 was 19-20% aa identity (41-43% nt) between CPs and 27-30% aa identity (41-46% nt) between RdRps.

3.4.2. Pterostylis sanguinea totivirus A (PsTVA): three isolates of a proposed new totivirus from orchid plants

Three related contigs of 4631 nt, 3631 nt and 3613 nt were identified from leaf tissue of P. sanguinea plants collected in 2013 (Table 1). The sequences resembled those of members of genus Totivirus. Totiviruses have dsRNA genomes consisting of a single molecule 4.6-7.0 kbp in length that encodes two, usually overlapping, ORFs (Wickner et al., 2012). The three putative totivirus isolates encoded by these three partial genome sequences were designated Pterostylis sanguinea totivirus A (PsTVA)
isolates Murdoch-8, Murdoch-9 and Murdoch-10 (GenBank accessions KU291927, KU291926 and KU291928 respectively). 2970 raw sequence reads were mapped to PsTVA-Murdoch-8 with pairwise identity of 81.7% and 61.9-fold mean coverage per base across its genome. Isolate Murdoch-9 was assembled from 94,091 reads at pairwise identity of 81.9% and mean coverage of 3142.8-fold per base. 68,653 reads were mapped to isolate Murdoch-10 with pairwise identity of 81.9% and mean coverage per base of 3153.0-fold. Each of the three sequences encoded two non-overlapping partial ORFs representing a CP and RdRp (Fig 1F), with no evidence of a slippery sequence indicating ribosomal frameshifting. The CPs had a L-A CP domain, which is typical of the yeast-infesting Totivirus type species Saccharomyces cerevisiae L-A virus (ScV-L-A). The L-A CP domain was located at nt residues 40-1212 (Murdoch-8), 1-245 (Murdoch-9) and 9-1142 (Murdoch-10). RdRp-4-like domains were detected at nt 3000-4088 (Murdoch-8), nt 2009-3097 (Murdoch-9) and nt 2948-3589 (Murdoch-10). Conserved RdRp core motifs V and VI (Koonin, 1991) were located on the genomes of isolate Murdoch-8 at aa residues 527-564 (nt 3774-3887) and isolate Murdoch-9 at aa 403-440 (nt 2783-2892) as SGx3 Tx3 NTx24 GDD. Comparison of the isolates showed 42-60% nt identity between genomes, 36-44% aa identity (48-56% nt) between CPs and 62-65% aa identity (61-64% nt) between RdRps. These figures are slightly below or above the suggested species demarcation limit for totivirus species of <50% aa identity (Wickner et al., 2012), indicating they may be categorised as the same species.

Blastp analysis of the two proteins revealed that the closest matches to described species were to the CP (32-37% aa identity) and RdRp (56-58% aa identity) of black raspberry virus F (BRVF; GenBank accession NC_009890), a proposed
totivirus described from leaf tissue or fungus-infected leaf tissue of *Rubus occidentalis* in the USA. PsTVA also grouped with totiviruses that infect fungi and yeast (Fig 2D).

3.5. Other virus-like sequences

In addition to the viruses described above, there is evidence from 30 short sequence fragments that were 454-3118 nt in length that other viruses were also present (Table S1). These are not described in detail here because they were estimated to represent less than 50% of genomes, thereby making assignment to taxonomic groups speculative. Analysis using Blastp revealed they most closely matched viruses from five virus families and five genera, and some were unclassified (Table S1). From leaf samples, a megabirnavirus-like sequence (P-2012), four other totivirus-like sequences (P-2013) and two related to viruses of the family *Amalgaviridae* (P-2013) were identified (Table S1). The totivirus-like contigs were closely related to the three PsTVA isolates and may represent more isolates of PsTVA, or belong to related species. They shared 41-95% nt identity between genomes, 44-95% nt (12-93% aa) identity between RdRp sequences and 48-57% nt (11-49% aa) identity between CP sequences.

From the two fungal isolates, a further 23 virus-like contigs were identified (454-3118 nt) that most closely resembled species within the genera *Endornavirus*, *Hypovirus*, *Megabirnavirus* and unclassified mycoviruses (Table S1). The four endornavirus-like contigs (454-1245 nt) detected in *Ceratobasidium* sp. (F-2013) were matched to endornaviruses recently identified from mycorrhizal fungi of other terrestrial orchids in the region (Ong et al., 2016b). RT-PCR was done using primers
specific to Ceratobasidium endornaviruses A-H, and they confirmed the presence of Ceratobasidium endornaviruses G and H (data not shown) (Ong et al., 2016b).

4. Discussion

A small wild population of three and four *P. sanguinea* plants collected in 2012 and 2013 and mycorrhizal fungi associated with two of the plants were found to be colonised by numerous persistent viruses, none of which had been described previously. In addition to the six viruses described above, there were alphapartitiviruses, betapartitiviruses and partitivirus-like sequences (family *Partitiviridae*) associated with the same mycorrhizal fungi (Ong et al., 2017). There were at least 10 partitiviruses – seven alphapartitiviruses and three betapartitiviruses – found in fungal isolate F-2012. From isolate F-2013, five alphapartitiviruses and one betapartitiviruses initially identified by Illumina sequence analysis. The majority of these partitiviruses were subsequently detected in both mycorrhizal strains using virus-specific primers (Ong et al., 2017). In total, at least 22 definitive viruses, proposed as belonging to the genera *Alphapartitivirus*, *Betapartitivirus*, *Hypovirus*, *Mitovirus*, *Totivirus* and unclassified mycoviruses were identified (Ong et al., 2017). All but two of the new viruses were identified from pure cultures of *Ceratobasidium* derived from pelotons isolated from two *P. sanguinea* plants. The findings extend the geographical range of probable members of *Betapartitivirus* (*Ceratobasidium* partitiviruses; Ong et al., 2017), *Hypovirus* (CbHV1; Ceratobasidium hypovirus-like contig 1-5), *Megabirnavirus* (Ceratobasidium megabirnavirus-like contigs 1-2; *Pterostylis* megabirnavirus-like contig 1), *Mitovirus* (CbMV1) and *Totivirus* (PsTV1; *Ceratobasidium* totivirus-like contigs 1-4), which had not previously been identified from Australia. Although tentatively assigned classification by phylogenetic analysis
with members of known groups, the proposed classifications are by no means certain because many sequences represented partial genomes, and most were genetically distant from described species.

4.1. Classification of new viruses

Most of the new viruses were tentatively classified with existing higher order taxa, but assigning them to existing lower order taxa was often problematical. For example, CbMV1 was proposed as a member of *Mitovirus*, but it does not fit easily into the two proposed subgroups within the genus (Doherty *et al.*, 2006). Instead it groups with other unclassified mycorrhizae-derived mitoviruses (Fig 2A) that usually encode tryptophan with UGG rather than UGA, which confers on them the capability of replicating in the host cytoplasm as well as its mitochondria (Kitahara *et al.*, 2014). Similarly, the proposed hypovirus CbHV1 is phylogenetically closest to the hypoviruses, but features of its genome organisation and host type (basidiomycete-infecting) place it outside the existing two hypovirus subgroups (Yaegashi *et al.*, 2012). Mycovirus-like PsVA, detected in leaf tissue, was grouped with unclassified mycoviruses in a new proposed genus, Phlegivirus. The two mycoviruses, CbV1 and CbV2 share sequence identity and genome organisation with other mycoviruses from different continents, but none are currently assigned taxa. Together, these findings indicate that the evolutionary history of mycoviruses is more complex than currently recognised.

PsTVA is the only new virus that shared a relatively close evolutionary history with a previously described virus – the proposed totivirus black raspberry virus F. Sequence identities of these two viruses are marginally above the species demarcation
threshold of 50% aa identity for totiviruses set by the ICTV (Wickner et al., 2012), but given that the host species are distinct and their locations are widely separated, we propose that they belong to different species.

4.2. Host identification

Viruses identified from plant materials are usually assumed to be plant viruses, with no distinction being made between viruses capable of replicating in plant cells or fungal cells that co-occur with plant tissues. It can be difficult to ascertain the true host. The current experiment was designed to detect RNA viruses that have dsRNA genomes, and ssRNA viruses that have a dsRNA replicative form. An alternative approach would be to sequence total leaf RNA (but depleted of ribosomal RNA) to eliminate the bias towards detection of dsRNA viral genomes and provide a clearer representation of the entire virome, including the presence of fungal transcripts. This metatranscriptomic approach was used recently in a study by Marzano et al. (2015), which identified 22 putative mycoviruses (dsRNA, ssRNA and ssDNA) from soybean leaf samples.

4.3. Viruses, fungi and orchids

In orchid biology, the plant-fungus symbiotic partnership is critical, but the roles viruses may play in this relationship remain largely unknown. Most mycoviruses appear to have little influence on fungal pathogenicity (Seo et al., 2004; Vainio et al., 2012), but some are demonstrated to influence their hosts. The most well-known example of a mycovirus that reduces fungal pathogenicity is CHV1 of Cryphonectria parasitica, the fungal pathogen that causes chestnut blight (Anagnostakis & Day, 1979). The term ‘Rhizoctonia decline’ was used to describe the decreased growth rate
and lack of sclerotia production of an isolate of *Rhizoctonia solani* infected with a dsRNA virus (Castanho & Butler, 1978a; Castanho and Butler, 1978b). In contrast, the plant pathogenic fungus *Nectria radicicola* became more virulent in the presence of a 6.0 kbp dsRNA virus (Ahn & Lee, 2001).

Ecological roles of mycoviruses may be elucidated when fungal cultures are cured of persistent viruses so that isogenic lines with and without mycoviruses can be compared. It has been reported that *in vitro* culturing of some fungi can cause them to lose mycoviruses (Bao & Roossinck, 2013; Roossinck, 2015). Because the *Ceratobasidium* isolates analysed here were cultured *in vitro*, this study may provide an underestimate of the number of mycoviruses that exist under natural conditions.

Ecological factors, such as drought, fire, grazing, low seed set, salinity of habitats and weed invasion, are known to impact negatively on orchid populations, and this is the case with many of the threatened orchid species in Western Australia (Coates & Atkins, 2001; Swarts & Dixon, 2009; Brundrett, 2016). Whether viruses play positive or negative roles in orchid biology remains unclear. The identification of viruses is a step in on-going studies of the interplay between wild plants, fungi and viruses. Such studies may shed light on understanding why populations of many orchids in south-west Australia (and globally) are shrinking alarmingly, while other orchid species are thriving to the point of becoming weeds, e.g. *Microtis media* and *Disa bracteata* (Bonnardeaux et al., 2007; Swarts & Dixon, 2009; De Long et al., 2013). More broadly, it may also provide clues to understanding the roles, positive or negative, that mycoviruses play in fungal interactions in other natural systems, and even in artificial systems such as agriculture.
Acknowledgement

This work was supported by Australian Research Council Linkage [Grant LP110200180] in collaboration with The Botanic Gardens and Parks Authority and The Australian Orchid Foundation.
References


Khalifa


Marchler-Bauer, A, C Zheng, F Chitsaz, MK Derbyshire, LY Geer, RC Geer, NR Gonzales, M Gwadz, DI Hurwitz, CJ Lanczycki, F Lu, S Lu, GH Marchler, JS Song,


Ong, JWL, H Li, K Sivasithamparam, KW Dixon, MGK Jones and SJ Wylie. 2016b. Novel Endorna-like viruses, including three with two open reading frames, challenge the membership criteria and taxonomy of the *Endornaviridae.* *Virology* 499: 203-211.

Ong, JWL, H Li, K Sivasithamparam, KW Dixon, MGK Jones and SJ Wylie. 2017. The challenges of using high-throughput sequencing to track multiple new bipartite


**Figures legends**

**Figure 1.** Proposed genome organisations of (A) Ceratobasidium mitovirus 1 (*Mitovirus*) (B) Ceratobasidium virus 1 (unclassified mycovirus) (C) Ceratobasidium virus 2 (unclassified mycovirus) (D) Ceratobasidium hypovirus 1 (*Hypovirus*) (E) Pterostylis sanguinea leaf-associated virus A (unclassified mycovirus-like) and (F) Pterostylis sanguinea leaf-associated totivirus A (*Totivirus*). Asterisks indicate incomplete 5' and/or 3' ends. Shaded boxes represent Nudix hydrolase (N) and RNA dependent RNA polymerase (RdRp) domains.

**Figure 2.** Maximum likelihood phylogenetic trees constructed from polyprotein or RdRp deduced amino acid sequences of (A) proposed mitovirus Ceratobasidium mitovirus 1 (indicated by a dot), (B) proposed Ceratobasidium virus 1 (indicated by dots), Ceratobasidium virus 2 (indicated by a square) and Pterostylis sanguinea leaf-associated virus A (indicated by a triangle), (C) proposed hypovirus Ceratobasidium hypovirus 1 (indicated by a dot), and (D) proposed totivirus Pterostylis sanguinea leaf-associated totivirus A (indicated by dots) with those of most closely related described viruses. 1000 bootstrap replications were carried out and branch confidence values below 60% were omitted. Branch lengths represent calculated evolutionary distance in units of amino acid substitutions per site.
Table legend

Table 1. Viruses (other than partitiviruses) identified from *Pterostylis sanguinea* orchids and associated *Ceratobasidium* sp. mycorrhizal fungi. Viruses identified from pooled leaf tissue of three and four *P. sanguinea* plants (P-2012, P-2013, respectively) are shown, and from *Ceratobasidium* isolates F-2012 and F-2013, each from the underground stem of one *P. sanguinea* plant.
Table 1. Viruses (other than partitiviruses) identified from *Pterostylis sanguinea* orchids and associated *Ceratobasidium* sp. mycorrhizal fungi.

Viruses identified from pooled leaf tissue of three and four *P. sanguinea* plants (P-2012, P-2013, respectively) are shown, and from *Ceratobasidium* isolates F-2012 and F-2013, each from the underground stem of one *P. sanguinea* plant.

<table>
<thead>
<tr>
<th>Proposed virus name</th>
<th>Isolate name</th>
<th>Proposed classification Family, Genus</th>
<th>Sequence length (nt) [protein(s) length (aa)]</th>
<th>Virus host (Sample no.)</th>
<th>Closest Blastp match</th>
<th>GenBank accession (e-value, % aa identity) of closest match</th>
<th>Estimated % ORF*</th>
<th>Estimated % genome*</th>
<th>GenBank accession</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ceratobasidium mitovirus 1 (CbMV1)</td>
<td>Murdoch-1</td>
<td>Narnaviridae, Mitovirus</td>
<td>2850 [817]</td>
<td>Ceratobasidium sp. (F-2012)</td>
<td>Rhizogagus sp. HR1 mitovirus like ssRNA</td>
<td>BAN85985 (9e-83, 31%)</td>
<td>100%</td>
<td>100%</td>
<td>KU291923</td>
</tr>
<tr>
<td>Ceratobasidium virus 1 (CbV1)</td>
<td>Murdoch-2</td>
<td>Unclassified</td>
<td>8227 [988, &gt;1593]</td>
<td>Ceratobasidium sp. (F-2012)</td>
<td>Desulfovibrio oxyclinae transposase (ORF1); Rosellinia necatrix mycovirus 1-W1032/S5 (ORF2)</td>
<td>WP_026167673 (2.2, 26%)</td>
<td>&gt;60%</td>
<td>&gt;90%</td>
<td>KU291947</td>
</tr>
<tr>
<td>Ceratobasidium virus 1 (CbV1)</td>
<td>Murdoch-3</td>
<td>Unclassified</td>
<td>7161 [1593, &gt;632]</td>
<td>Ceratobasidium sp. (F-2012)</td>
<td>Desulfovibrio oxyclinae transposase (ORF1); Rosellinia necatrix mycovirus 1-W1032/S5 (ORF2)</td>
<td>WP_026167673 (0.75, 26%)</td>
<td>100%</td>
<td>&gt;50%</td>
<td>KU291948</td>
</tr>
<tr>
<td>Ceratobasidium virus 1 (CbV1)</td>
<td>Murdoch-4</td>
<td>Unclassified</td>
<td>4051 [&gt;1123]</td>
<td>Ceratobasidium sp. (F-2012)</td>
<td>Desulfovibrio oxyclinae transposase (ORF1); Rosellinia necatrix mycovirus 1-W1032/S5 (ORF2)</td>
<td>WP_026167673 (0.75, 26%)</td>
<td>&gt;70%</td>
<td>45%</td>
<td>KU291949</td>
</tr>
<tr>
<td>Ceratobasidium virus 2 (CbV2)</td>
<td>Murdoch-5</td>
<td>Unclassified</td>
<td>7089 [&gt;1112, &gt;1162]</td>
<td>Ceratobasidium sp. (F-2012)</td>
<td>Rhizoctonia solani RNA virus HN008</td>
<td>NP_009158859 (2e-04, 25%)</td>
<td>&gt;90%</td>
<td>&gt;90%</td>
<td>KU291938</td>
</tr>
<tr>
<td>Ceratobasidium hypovirus 1 (CbHV1)</td>
<td>Murdoch-6</td>
<td>Hypoviridae, Hypovirus</td>
<td>7143 [1211, &gt;1085]</td>
<td>Ceratobasidium sp. (F-2012)</td>
<td>Cryphonectria hypovirus 1</td>
<td>NP_041091 (1e-12, 30%)</td>
<td>100%</td>
<td>&gt;30%</td>
<td>95%</td>
</tr>
<tr>
<td>Pterostylis sanguinea leaf-associated virus A (PsVA)</td>
<td>Murdoch-7</td>
<td>Unclassified</td>
<td>10,716 [1912, &gt;1252]</td>
<td><em>P. sanguinea</em> (P-2012)</td>
<td>Lentinula edodes mycovirus HKA</td>
<td>BAM34027 (1e-105, 26%)</td>
<td>100%</td>
<td>&gt;90%</td>
<td>KU291925</td>
</tr>
</tbody>
</table>
Pterostylis sanguinea leaf-associated totivirus A (PsTVA)

Murdoch-8 Totiviridae, Totivirus 4631 [>723, >812] P. sanguinea (P-2013) Black raspberry virus F YP_001497150 (6e-139, 36%) >90% 92% KU291927

Murdoch-9 Totiviridae, Totivirus 3631 [>382, >685] P. sanguinea (P-2013) Black raspberry virus F YP_001497150 (2e-52, 32%) >50% 72% KU291926

Murdoch-10 Totiviridae, Totivirus 3613 [>696, >404] P. sanguinea (P-2013) Black raspberry virus F YP_001497150 (2e-152, 37%) >90% 72% KU291928

\(^*\)Calculation of genome and protein percentage was based on sequence length of closest blastp match and/or related virus isolates.

\(^*\)Estimated percentage of protein was limited by lack of complete ORF