Diagnosis and discovery of fungal viruses using deep sequencing of small RNAs

Eeva J. Vainio,1 Jaana Jurvansuu,1 Janne Streng,2 Minna-Liisa Rajamäki,2 Jarkko Hantula1 and Jari P. T. Valkonen2

Correspondence
Eeva J. Vainio
eeva.vainio@metla.fi

1Finnish Forest Research Institute, Vantaa Research Unit, Jokiniemenkuja 1, Vantaa, Finland
2Department of Agricultural Sciences, P.O. Box 27, University of Helsinki, FI-00014 Helsinki, Finland

Analysis of virus-derived small RNAs with high-throughput sequencing has been successful for detecting novel viruses in plants and invertebrates. However, the applicability of this method has not been demonstrated in fungi, although fungi were among the first organisms reported to utilize RNA silencing. Here, we used virus-infected isolates of the fungal species complex Heterobasidion annosum sensu lato as a model system to test whether mycovirus genome segments can be detected with small RNA deep sequencing. Species of the genus Heterobasidion are some of the most devastating forest pathogens in boreal forests. These fungi cause wood decay and are commonly infected with species of the family Partitiviridae and the unassigned virus species Heterobasidion RNA virus 6. Small RNA deep sequencing allowed the simultaneous detection of all eight double-stranded RNA virus strains known to be present in the tested samples and one putative mitovirus species (family Narnaviridae) with a single-stranded RNA genome, designated here as Heterobasidion mitovirus 1. Prior to this study, no members of the family Narnaviridae had been described as infecting species of Heterobasidion. Quantification of viral double- and single-stranded RNA with quantitative PCR indicated that co-infecting viral species and viruses with segmented genomes can be detected with small RNA deep sequencing despite vast differences in the amount of RNA. This is the first study demonstrating the usefulness of this method for detecting fungal viruses. Moreover, the results suggest that viral genomes are processed into small RNAs by different species of Heterobasidion.

INTRODUCTION

RNA interference (RNAi), also called RNA silencing, refers to post-transcriptional control of gene expression and protection of organisms against mobile repetitive elements, retrotransposons and viral infections (Tijsterman et al., 2002). The cellular machinery for RNAi is present in animals, plants and fungi (Meister & Tuschl 2004; Chang et al., 2012). During RNA silencing, Rnase III endoribonucleases called Dicers (or Dicer-like in plants) detect and cleave double-stranded RNA (dsRNA) templates into small interfering RNAs (siRNAs), which are typically 21, 22 or 24 nt long in plants (Bologna & Voinnet, 2014). The siRNAs then associate with an RNA-induced silencing complex in which a protein of the Argonaute family catalyses the sequence-guided degradation of target mRNA. Because the RNAi machinery targets potentially deleterious non-self nucleic acids, virus-infected host organisms are typically enriched for siRNAs for viral sequences.

Consequently, the analysis of virus-derived small RNAs with high-throughput sequencing has been successfully used in plants and invertebrates to detect novel viruses (Kreuze et al., 2009; Pantaleo et al., 2010; Wu et al., 2010; Bi et al., 2012; Kashif et al., 2012).

Among fungi, RNAi was first observed in the ascomycetous Neurospora crassa (Romano & Macino, 1992); the transformation of N. crassa with multiple adjacent copies of genes required for carotenoid pigment biosynthesis unexpectedly results in reduced pigmentation, a phenomenon the authors called ‘quelling’. Segers et al. (2007) demonstrated for the first time that RNAi acts as an antiviral defence mechanism in fungi by showing that Dicer gene disruptions in the ascomycetous Chestnut blight fungus, Cryphonectria parasitica, result in increased viral RNA levels. The loss of RNAi in the budding yeast, Saccharomyces cerevisiae (phylum Ascomycota), is believed to result from adaptation to the presence of Saccharomyces cerevisiae L-A virus. This virus and its satellite-encoded ‘killer’ toxins provide a selective advantage for the host in competition against yeast strains lacking immunity to the toxin (Drinnenberg et al., 2009, 2011).

The GenBank/EMBL/DDBJ accession number for the complete genome sequence of HetMV1-an1 is KJ873059.

Three tables are available with the online Supplementary Material.
The *Heterobasidion annosum sensu lato* species complex includes some of the most devastating forest pathogens in boreal forests. These fungi spread efficiently via basidiospores that infect fresh stump surfaces and butt wounds of living trees (Stenlid & Redfern, 1998). They also spread vegetatively via root contacts and single clonal individuals can survive for decades and infect dozens of trees. According to their host tree preferences, and genetic evidence, species of *H. annosum s. lat.* can be divided into two major clades: the 'spruce clade' includes the Eurasian species *Heterobasidion parviporum* and the Southern European *Heterobasidion abietinum* together with the North American *Heterobasidion occidentale*, whereas the European *Heterobasidion annosum* and the North American *Heterobasidion irregularare* form the ‘pine clade’. The recently published genome of *H. irregularare* contains the core components of the RNAi machinery, such as three Dicer-encoding genes and seven Argonaute-encoding genes (Olson et al., 2012; Hu et al., 2013).

dsRNA virus infections occur in approximately 15–17% of cultured *Heterobasidion* strains (Ihrmark, 2001; Vainio et al., 2011b). Most of these infections do not cause drastic phenotypic effects on their hosts when cultured in the laboratory, but some affect the growth of the host fungus (Vainio et al., 2010, 2012) or may reduce the germination frequency of basidiospores (Ihrmark et al., 2004). Recently, we showed that a single *Heterobasidion*-infecting virus strain may have beneficial, commensal or detrimental effects on a single host isolate depending on culture conditions and interacting antagonistic fungi (Hyder et al., 2013). Viruses that infect *Heterobasidion* species are transmitted between host strains during anastomosis of fungal hyphae, which also allows virus transmission between incompatible fungal strains and distantly related species of *Heterobasidion* (Ihrmark, 2001; Vainio et al., 2010, 2011a, b). Known viruses of *Heterobasidion* also spread via basidiospores and conidia (Ihrmark et al., 2002, 2004).

*Heterobasidion* RNA virus 6 (HetRV6) is responsible for ~70% of the dsRNA virus infections among *Heterobasidion* strains in Europe (Vainio et al., 2012). The RNA-dependent RNA polymerase (RdRp) gene of HetRV6 shares a moderate level of sequence similarity with the *Curvularia* thermal tolerance virus (CThTV), which is involved in a three-way symbiosis with its host fungus (the endophytic ascomycete *Curvularia protuberata*) and a grass species (Márquez et al., 2007). Whether HetRV6 encompasses a second genomic segment similar to CThTV is unknown. Members of the virus family *Partitiviridae* infect all five species of *H. annosum s. lat.* (Ihrmark, 2001; Vainio et al., 2010, 2011a, b, 2013b). Partitiviruses have bisegmented dsRNA genomes that encode an RdRp and a capsid protein (CP). The two genomic RNAs are separately encapsidated (King et al., 2012; Nibert et al., 2013). The RdRp of partitiviruses mediates plus-strand RNA synthesis (transcription) inside the virus particle using the minus strand of the genomic dsRNA as a template, after which the genomic plus strand is extruded for use in translation in the host cytosol (Nibert et al., 2014). Mitochondrial viruses (genus *Mitovirus*) of the family *Narnaviridae* are common in fungi (Hillman & Cai, 2013), but only two mitovirus species have been described in basidiomycete fungi and no mitoviruses are known to infect *Heterobasidion* spp. Fungal mitoviruses are unencapsidated viruses with positive-sense single-stranded RNA (ssRNA) genomes usually predicted to contain only a single gene that encodes RdRp (King et al., 2012).

*Cryphonectria parasitica* strains with Dicer gene disruptions have been used to demonstrate that RNAi targets the hypoviruses and mycoreoviruses of *Cryphonectria parasitica* (Segers et al., 2007; Zhang et al., 2008, 2012) as well as a partitivirus and a victorivirus that naturally infect the ascomycetous root pathogen *Rosellinia necatrix* (Chiba et al., 2013a, b). Hypoviruses have positive-sense ssRNA genomes and cause hypovirulence in *Cryphonectria parasitica*, whereas mycoreoviruses, partitiviruses and victoriviruses have dsRNA genomes (King et al., 2012). To our knowledge, RNAi against the HetRV6/CThTV-like virus clade or mitoviruses has not been demonstrated. Our aim was to test the following hypotheses: (i) small RNA deep sequencing can be used for virus detection in fungi, (ii) *Heterobasidion* species process viral RNAs into siRNAs and (iii) small RNA deep sequencing can be used to discover previously unknown viral species from *Heterobasidion* spp.

**RESULTS**

**Yield and size distribution of the siRNA reads**

The small RNA deep sequencing analysis included small RNA from ten fungal strains representing four different species of *Heterobasidion*. Seven of these strains were known to be infected with species of the family *Partitiviridae* or HetRV6 (Table 1). The total number of reads produced by the Illumina sequencer was 27014983 (read length 10–34 nt), corresponding to 1351 megabases. Most clusters (92%) fulfilled the default Illumina quality criteria and 96% of bases had a quality score ≥30. A relatively high proportion (49%) of the reads consisted of RNAs of 10–17 nt, whereas reads of 18–26 nt accounted for 43% of the RNAs. The two most prominent read lengths were 13 nt and 19 nt (Fig. 1).

**Sequence contigs produced with Velvet**

The Velvet program was used for *de novo* assembly of contigs using the small RNA reads. Contig assembly was first conducted using reads of 21–24 nt, which are typical lengths for siRNAs generated by many Dicer proteins (Chang et al., 2012; Bologna & Voinnet, 2014). However, because a considerable proportion of reads were shorter than 21 nt (Fig. 1), we also examined the presence of viral sequences among the shorter siRNAs. Viral contigs were obtained with Velvet using reads of 18 nt and larger, and therefore we included all reads of 18–27 nt in the assembly of the reads.
subsequent Velvet analyses. Using k-mer values smaller than 13 or larger than 19 resulted in significant decreases in contig number (the k-mer value indicates the word size used in the de Bruijn graphs) (Zerbino & Birney, 2008). We selected a k-mer value of 13 to be used in an optimized Velvet run in which host rRNA sequences were subtracted from the sequence pool using the Coderestar script and the primary contigs were then post-assembled with the AssemblyAssembler program into larger contigs. This optimized Velvet run resulted in 1608 contigs 56–515 nt in length and showed that viral contigs accounted for 3.7% (n=60) of all contigs detected. Using k-mer values of 15, 17 and 19, the number of contigs produced was 1421, 747 and 114, and the largest contigs were 544, 765 and 584 nt in size, respectively. All Velvet assemblies that produced any contigs were examined for the presence of viral sequences using BLASTN and BLASTX.

Identification of viral sequences with BLAST

The contigs produced with Velvet were used to search similar sequences in the NCBI/GenBank database at the nucleotide and protein level using BLASTN and BLASTX. This analysis revealed HetPV2-pa1 and HetPV4 sequences encoding RdRp and CP; and HetPV4-pa1, HetPV13-an1 and HetRV6-pa19-c sequences encoding RdRp. Heterobasidion partitivirus 2 (HetPV2) and HetPV7 were recently classified by the International Committee on the Taxonomy of Viruses into the genus Betapartitivirus (Nibert et al., 2014) and share 65.5% polymerase sequence identity at the nucleotide level. The remaining partitiviruses included in the analysis represent the genus Alphapartitivirus and share 39.7–65.6% polymerase sequence identity at the nucleotide level (Table S1, available in the online Supplementary Material). In addition to these viruses that are known to infect the fungal strains used in our study, we found some contigs that resembled fungal mitoviruses. In the optimized Velvet run, we found nine mitoviral contigs 58–241 nt in length that had 42–94% sequence identity with known mitoviruses, as determined using BLASTX.

Table 1. *Heterobasidion* isolates used for small RNA deep sequencing

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain</th>
<th>Virus infection(s)*</th>
<th>GenBank</th>
<th>Taxonomy</th>
<th>Reference†</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>H. abietinum</em></td>
<td>04188</td>
<td>HetRV6-ab16</td>
<td>HQ189459</td>
<td>Unassigned</td>
<td>Vainio et al., (2012)</td>
</tr>
<tr>
<td><em>H. abietinum</em></td>
<td>RF2s2</td>
<td>Not known</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><em>H. annosum</em></td>
<td>05021</td>
<td>Not known</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td><em>H. irregulare</em></td>
<td>704</td>
<td>HetPV17-ir1</td>
<td>KJ873060</td>
<td>Alphapartitivirus</td>
<td>E. J. Vainio and others, unpublished</td>
</tr>
<tr>
<td><em>H. parviporum</em></td>
<td>7R18</td>
<td>HetPV2-pa1</td>
<td>HM565953 (RdRp)</td>
<td>Betapartitivirus</td>
<td>Vainio et al., (2011a)</td>
</tr>
<tr>
<td><em>H. parviporum</em></td>
<td>7R226</td>
<td>HetPV7-pa1</td>
<td>JN606091 (RdRp)</td>
<td>Betapartitivirus</td>
<td>Vainio et al., (2014)</td>
</tr>
<tr>
<td><em>H. parviporum</em></td>
<td>RT3.49C</td>
<td>HetPV4-pa19-c</td>
<td>KF551892</td>
<td>Unassigned</td>
<td>Vainio et al., (2011b, 2013a)</td>
</tr>
<tr>
<td><em>H. parviporum</em></td>
<td>10084</td>
<td>Not known</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

*Note that the nomenclature of partitiviruses has recently changed. HetPV2 and HetPV4 were formerly designated HetRV2 and HetRV4, respectively (Nibert et al., 2014).†All other virus strains except HetMV1-an1 were known to be present in the tested fungal strains, whereas HetMV1-an1 was newly identified with small RNA deep sequencing in this study.

Fig. 1. Size distribution of small RNA reads obtained using Illumina deep sequencing.
Mitovirus sequence characterization and phylogenetic analysis

The short (58–241 nt) Velvet contigs with sequence similarity to mitoviruses were assembled using previously described mitoviral genome sequences at the nucleotide and protein levels to reveal their putative genomic locations and to design PCR primers for virus screening. The cDNAs of the ten *Heterobasidion* strains used in this study were screened one at a time using mitovirus primers and revealed the presence of a new putative mitovirus, designated here as HetMV1-an1 (*Heterobasidion* mitovirus 1, strain 1 from *H. annosum*), in *H. annosum* 94233. Subsequently, the intervening sequence regions that were not elucidated with small RNA analysis were amplified with PCR using specific primers and the products were sequenced. The resulting complete genome sequence of HetMV1-an1 (4379 nt in length) was deposited in GenBank (accession number KJ873059). Mitoviral amplification products were not detected using total cellular nucleic acids (including nuclear and mitochondrial DNA) as a PCR template and therefore HetMV1-an1 appeared not to be incorporated into the host genome, in contrast to a mitovirus described in the basidiomycete fungus *Thanatephorus cucumeris* (anamorph *Rhizoctonia solani*; Table 2) (Lakshman et al., 1998).

The host of HetMV1-an1 (*H. annosum* 94233) is co-infected with the partitivirus HetPV13-an1 (GenBank accession numbers KF631777 and KF631778). However, we did not detect mitoviral dsRNA during genome characterization of the co-infecting partitivirus (Kashif et al., 2014). Mitoviruses are positive-sense ssRNA viruses, but their replicative intermediates occur in double-stranded form and they can often be detected as dsRNA in the presence of a sufficiently high copy number (Tuomivirta & Hantula 2003; Martínez-Alvarez et al., 2014).

Sequence comparisons (Table 2) and phylogenetic analysis (Fig. 2) indicated that HetMV1-an1 is a novel putative mitovirus species that shares approximately 33% identity at the protein level with its closest relative, *Clitocybe odora* mitovirus (Heinz, 2012). BLASTP comparisons revealed sequence identities of 32–48% with other mitoviral species (Table 2), but this analysis recognized only sequence regions that overlapped with the conserved mitovirus RdRp protein family domain (pfam05919; aa residues 339–652 in the predicted RdRp of HetMV1-an1). At the nucleotide level, BLASTN recognized only short nucleotide stretches of related mitoviruses, most covering less than 1% of the HetMV1-an1 genome (Table 2). The 4379 nt HetMV1-an1 sequence has a DNA G+C content of 40.0 mol%. The genome of HetMV1-an1 contains an ORF encoding RdRp of 836 aa when translated using mitochondrial codon usage (translation table 4). This ORF is interrupted by nine UGA stop codons when translated using the standard nuclear translation table (Fig. 3). No other long ORFs were detected. The lengths of the 5′- and 3′-untranslated regions were 801 nt and 1067 nt, respectively.

Guide strand-mediated assembly with the MAQ sequence mapping and assembly program

The MAQ program was used to map small RNA reads to reference sequences, namely all characterized viral genome segments found previously in the *Heterobasidion* strains used in our study (Table 1). Small RNAs covered 90.2% and 99.3% of the polymerase-encoding genome segments and 97.3% and 96.6% of the capsid-encoding segments of the two betapartitiviruses, HetPV2-pa1 and HetPV7-pa1, respectively. The average sequencing depth for these segments varied from 46x to 310x (Table 3). The polymerase-encoding segments of HetRV6-ab6, HetRV6-pa19-c, HetPV4-pa1, HetPV13-an1 and HetMV1-an1 were also well represented among small RNAs (90.2–98.0% sequence coverage and sequencing depth of 14–147x; Table 3). However, virus strains HetPV15-pa1 and HetPV17-ir1 were less well covered by small RNAs (sequence coverage 28.8–38.1%) and their MAQ scores were less reliable (Table 3). Moreover, sequencing depth was very low for these two viruses (Table 3) and they were not detected by Velvet. In all cases, the distal ends of the genome segments seemed to be poorly detected by MAQ. Nevertheless, all viruses known to be present in the sample were detected by MAQ.

The CP-encoding segment of HetPV13-an1 seemed to be less well covered by small RNAs than the RdRp gene (sequence coverage 76.4% and 96.0%, respectively). Moreover, the MAQ scores were less reliable for the CP segment compared with RdRp (Table 3), suggesting that the two genome segments may have differed with respect to sheer quantity. This idea was supported by quantitative PCR (qPCR) analysis, which showed that the RdRp:CP ratio in HetPV13-an1 was 12.0 for dsRNA and 8.7 for viral RNA (consisting mainly of ssRNA; Fig. 4). The opposite situation was observed for HetPV2-pa1, which seemed to produce more CP-derived small RNA than RdRp-derived small RNA (sequence coverage 97.3% and 90.2%, respectively).

The presence of two highly similar HetRV6 strains that share 92.5% sequence identity (HetRV6-ab6 and HetRV6-pa19-c) resulted in some ambiguity in the MAQ base calls (for example, sequence positions differing for C or T between the two sequences resulted in a Y base call), presumably owing to cross-assembly. On the other hand, clear signature nucleotides were observed that indicated the presence of both virus strains among the siRNAs (Table S2).

In the case of the partitiviruses, polymerase-encoding nucleotide sequence identity among strains included in the sample pool was moderate or low (39.7–65.6%; Table S1). The CP-encoding sequences of HetPV4-pa1 and HetPV17-ir1 have not been determined and whether HetRV6 has a CP is not known; we therefore attempted to identify these genome segments using CP gene sequences from the most closely related viruses (as determined with RdRp sequence comparisons using NCBI BLASTN) for MAQ runs (Table S3). However, the analysis yielded only low-quality sequence matches and, despite testing several PCR primers,
we obtained no amplification products of expected sizes (data not shown). This might have been expected as the CP-encoding segments of partitivirus species typically share lower sequence identity compared with the corresponding RdRp-encoding segments (Nibert et al., 2014), which shared 68–73 % sequence identity in our analysis (Table S3). Alternatively, the CP-encoding segments of HetPV4-pa1 and HetPV17-ir1 may not be subjected to RNA silencing.

### Amounts of viral RNAs in co-infected isolates

Two *Heterobasidion* isolates with viral co-infections were used to examine the relative amounts of viral dsRNA and ssRNA present in the host mycelia. The double-infected *H. parviporum* 7R226 isolate harbours the partitivirus HetPV7-pa1 and the unassigned HetRV6-pa19-c. The polymerase genes for HetRV6-pa19-c and HetPV7-pa1 were both well represented among the small RNAs (sequence coverage 94.6 % and 99.3 %, respectively). However, based on qPCR, 5.9 times more HetPV7-pa1 polymerase dsRNA was present compared with HetRV6-pa19-c polymerase dsRNA in the co-infected host isolate (Fig. 5). Notably, substantial differences were apparent in viral transcription activity, as 19.7 times more HetPV7-pa1 RdRp transcripts (i.e. viral RNA including mainly ssRNA) were present than HetRV6-pa19-c RdRp transcripts (Fig. 5).

*H. annosum* 94233 is co-infected with HetMV1-an1 and the partitivirus strain HetPV13-an1, which showed small RNA sequence coverages of 98 % and 96 %, respectively, for the polymerase gene. Based on qPCR, 2.3 times more mitoviral dsRNA was present compared with partitiviral dsRNA in the infected fungal strain. However, on average, 200 and 1780 times more HetMV1-an1 ssRNA than HetPV13-an1 RdRp and CP transcripts, respectively, were present (Fig. 4).

### DISCUSSION

This study demonstrates that small RNA deep sequencing can be used to detect fungal viruses and suggests that *Heterobasidion* species process viral RNA into small RNAs. The method allowed the simultaneous detection of nine mycoviral strains with dsRNA or ssRNA genomes, including a new putative mitovirus. The sequence of this previously unknown mitovirus shared only ~33 % sequence similarity with its closest relative, the Clitocybe odora mitovirus (Heinze, 2012), indicating that relatively divergent viral species can be detected with small RNA deep sequencing. However, as the identification of viral contigs is based on the availability of reference sequences in public databases, this method cannot be used to detect viruses with no known relatives. Hence, small RNA-based diagnostics of viruses benefit from the accumulation of sequence data of novel

<table>
<thead>
<tr>
<th>Virus species (RdRp)</th>
<th>Accessions (nt, aa)</th>
<th>Length (nt)</th>
<th>BLASTP identity/query cover (%)*</th>
<th>BLASTN identity/query cover (%)</th>
<th>Overall identity according to MAFFT†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clitocybe odora virus</td>
<td>HE717021.1, YP_005352912.1</td>
<td>3765</td>
<td>53/62</td>
<td>68/13</td>
<td>32.7 %</td>
</tr>
<tr>
<td>Ophiostoma mitovirus 3b</td>
<td>AM087550, CAJ32468.1</td>
<td>2332</td>
<td>45/42</td>
<td>NS</td>
<td>27.4 %</td>
</tr>
<tr>
<td>Botrytis cinerea mitovirus 1</td>
<td>EF808100, ABQ65153.3</td>
<td>2804</td>
<td>45/41</td>
<td>NS</td>
<td>26.7 %</td>
</tr>
<tr>
<td>Ophiostoma mitovirus 1b</td>
<td>AM087549, CAJ32467.1</td>
<td>2572</td>
<td>44/45</td>
<td>NS</td>
<td>25.2 %</td>
</tr>
<tr>
<td>Tuber aestivum mitovirus</td>
<td>HQ992989.1, YP_004564622.1</td>
<td>3480</td>
<td>44/50</td>
<td>NS</td>
<td>26.8 %</td>
</tr>
<tr>
<td>Ophiostoma mitovirus 3a</td>
<td>AJ004930.1, NP_660176.1</td>
<td>2617</td>
<td>48/42</td>
<td>NS</td>
<td>25.9 %</td>
</tr>
<tr>
<td>Sclerotinia homeocarpa mitovirus</td>
<td>AY172454, AAO21337.1</td>
<td>2632</td>
<td>47/46</td>
<td>NS</td>
<td>25.9 %</td>
</tr>
<tr>
<td>Sclerotinia sclerotiorum mitovirus 3</td>
<td>JX401537, AGC24232.1</td>
<td>2588</td>
<td>46/28</td>
<td>NS</td>
<td>23.6 %</td>
</tr>
<tr>
<td>Ophiostoma mitovirus 1a</td>
<td>AM087548, CAJ32466.1</td>
<td>3147</td>
<td>36/51</td>
<td>NS</td>
<td>24.4 %</td>
</tr>
<tr>
<td>Sclerotinia sclerotiorum mitovirus 6</td>
<td>KJ462513, AHX84133.1</td>
<td>2586</td>
<td>40/44</td>
<td>NS</td>
<td>24.9 %</td>
</tr>
<tr>
<td>Sclerotinia sclerotiorum mitovirus 14</td>
<td>KF913889, AHF48630.1</td>
<td>2564</td>
<td>40/44</td>
<td>NS</td>
<td>24.8 %</td>
</tr>
<tr>
<td>Thanatephorus cucumeris mitovirus</td>
<td>TCU51331, AAD17381.1</td>
<td>3570</td>
<td>32/58</td>
<td>NS</td>
<td>24.1 %</td>
</tr>
</tbody>
</table>

NS, No significant similarity (query cover less than 5 %).

*Query cover was identified with BLASTP and BLASTN and indicates the proportion (%) of the complete HetMV1-an1 sequence of 4379 nt (KJ873059) that was aligned with each particular reference sequence available in GenBank. The e-values for BLASTP comparisons were 7–76–5–44 and the BLASTN e-value for Clitocybe odora virus was 8–15.

†Overall sequence identity in a sequence alignment produced by MAFFT using the complete sequence for HetMV1-an1.

Table 2. Sequence similarity of HetMV1-an1 compared with related mitoviruses as determined with NCBI BLAST and sequence alignments
mycoviral species in public databases (Osaki et al., 2006; Chiba et al., 2009; Feldman et al., 2012; Lin et al., 2012; Yaegashi et al., 2013).

Recent next-generation sequencing studies have used purified dsRNA as a template and detected mycoviral sequences in grapevine plants (Coetzee et al., 2010; Al Rwahnih et al., 2011) and endophytic fungi (Feldman et al., 2012). Direct mRNA sequencing has revealed novel virus species in plants (Al Rwahnih et al., 2009; Wylie & Jones, 2011; Wylie et al., 2013) and fungi (Schoebel et al., 2014). Compared with mRNA sequencing, small RNA deep sequencing enriches the sequence pool for viral sequences and the advantage of this method over high-throughput sequencing of dsRNA is that small RNA deep sequencing may allow detection of a wider variety of viruses, including those with ssRNA or DNA genomes (Kreuze et al., 2009; Kashif et al., 2012). The novel putative mitovirus detected in this study has an ssRNA genome and remained undetected using a conventional cloning approach (Kashif et al., 2014). Among fungi, only one virus species with a DNA genome has been described (Yu et al., 2010), and whether the RNA transcripts of fungal DNA viruses are processed into small RNAs by the hosts remains unknown.

In this study, the sequence-guided assembly by MAQ revealed high small RNA coverage for most viruses known to be present in the tested fungal strains. The genome of the betapartitivirus HetPV7-pa1 was almost fully covered (sequence coverages of 99.3% and 96.6% for the RdRp- and CP-encoding segments, respectively). The genomes of HetRV6-ab6, HetRV6-pa19-c, HetPV4-pa1 and HetPV2-pa1 were also well represented among small RNAs (90.2–97.3% sequence coverage). However, HetPV15-pa1 and HetPV17-ir1 showed sequence coverages of only 28.8–38.1%. In addition to viral copy number, differences in the amount of small RNAs from different virus species may be related to the regulation of RNAi by the host or virus. Recent studies have shown that suppression of RNAi is a common phenomenon among fungal viruses. Hammond
Table 3. Coverage of small RNAs determined with the MAQ sequence mapping and assembly program

<table>
<thead>
<tr>
<th>Virus strain</th>
<th>RNA-dependent RNA polymerase (RdRp)</th>
<th>Capsid protein (CP)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reference sequence length (nt)</td>
<td>Coverage (%)*</td>
</tr>
<tr>
<td>HetPV2-pa1</td>
<td>2290</td>
<td>90.2</td>
</tr>
<tr>
<td>HetPV4-pa1</td>
<td>2007</td>
<td>93.6</td>
</tr>
<tr>
<td>HetPV7-pa1</td>
<td>2297</td>
<td>99.3</td>
</tr>
<tr>
<td>HetPV13-an1</td>
<td>1873</td>
<td>96.0</td>
</tr>
<tr>
<td>HetPV15-pa1</td>
<td>1882</td>
<td>28.8</td>
</tr>
<tr>
<td>HetPV17-ir1</td>
<td>2016</td>
<td>29.8</td>
</tr>
<tr>
<td>HetRV6-ab6</td>
<td>2050</td>
<td>91.6</td>
</tr>
<tr>
<td>HetRV6-pa19-c</td>
<td>1957</td>
<td>94.6</td>
</tr>
<tr>
<td>HetMV1-an1</td>
<td>4379</td>
<td>98.2</td>
</tr>
</tbody>
</table>

NA, Not available – the CP sequence has not been described.
ND, No data – a putative CP has not been detected for strains of HetRV6;
–, The mitovirus genome does not include a CP.

*Sequence coverage is the percentage sequence overlap determined by aligning the MAQ output sequences, indicating sites covered by siRNAs, with the original reference sequence.
†The proportion (%) of the mapped sequence with a quality score of ≥ 90 %.
‡Average sequencing depth across all non-gap regions as determined by MAQ.
et al. (2008) showed that RNAi in Aspergillus nidulans is suppressed by Aspergillus virus 1816 and Aspergillus virus 341 is targeted for degradation by RNAi by the same host. Similarly, Chiba et al. (2013b) showed that a victorivirus (family Totiviridae) of R. necatrix is targeted by RNAi and a mycoreovirus infecting the same host suppresses RNAi (Yaegashi et al., 2013). The siRNA read length profile observed in this study differed from a characteristic profile produced by Dicer-like RNases and a significant proportion of the small RNAs were shorter than the typical small RNAs of 21–25 nt produced by the ascomycetes N. crassa and Cryphonectria parasitica or the zygomycete Mucor circinelloides (Chang et al., 2012). The functionality of the three putative Dicer-encoding genes and seven Argonaute-encoding genes identified in the genome of H. irregulare has not been demonstrated experimentally. In Cryphonectria parasitica, the most prominent size class of hypovirus-derived siRNAs is 20–22 nt, but viral siRNAs are also present among siRNAs of 19 nt and 23–28 nt (Zhang et al., 2008).

In this study, two of the Heterobasidion host strains were co-infected with distantly related virus species. Quantification of viral dsRNA and ssRNA indicated that co-infecting viral species can be detected with small RNA deep sequencing despite vast differences in the amount of RNA. In the case of HetMV1/HetPV13 co-infection, we found over 200 times more mitovirus RNA than partitivirus RdRp RNA, although the ratio of HetMV1 : HetPV13 dsRNA was only 2.3. The mitoviral coding strand is present in infected tissue in a greater molar amount than the dsRNA, which is considered to be the replicative form of these positive-sense ssRNA viruses (King et al., 2012). The HetMV1-an1 RdRp-encoding sequence contains nine internal stop codons when translated using the standard nuclear translation table. Therefore, the mitoviral coding strand is predicted to be translatable only in mitochondria by mitochondrial codon usage. In the case of HetRV6/HetPV7 co-infection, we also found a significant difference in the amount of viral polymerase transcripts (ssRNA) and genomic RNA from the two co-infecting viruses. Whether co-infecting Heterobasidion viruses have additive, synergistic or antagonistic interactions in the co-infected hosts remains unclear. Co-infection may affect genome copy numbers, as has been demonstrated among many plant pathogenic viruses (Poolpol & Inouye, 1986; Mukasa et al., 2006).

![Graph](http://vir.sgmjournals.org)
We observed differences in the siRNA accumulation between the two genome segments of the partitiviruses, suggesting that the quantity of the two RNAs differs or that they are differentially targeted by RNAi. In particular, the CP-encoding RNA segment of HetPV13-an1 appeared to be less well represented in the small RNA pool compared with the RdRp segment. This quantitative discrepancy between HetPV13-an1 genome segments was supported by qPCR analysis. Moreover, Juvansuu et al. (2014) found that the ratio of RdRp- and CP-encoding dsRNA for HetPV2-pa1 is on average 0.55, which is in accordance with the results of our study, suggesting that CP of HetPV2-pa1 is better represented in the siRNA pool compared with RdRp.

Taken together, this study shows that *Heterobasidion* species are able to process the RNA of various viruses into small RNAs; in addition to partitiviruses these viruses include the unassigned taxon HetrV6 and the new putative species of the genus *Mitovirus*. This observation is noteworthy as the entire reproductive cycle of mitoviruses is thought to occur inside host mitochondria (King et al., 2012; Hillman & Cai, 2013). *Thanatephorus cucumeris* and *Heterobasidion* are able to process the RNA of various viruses into small RNAs; in addition to partitiviruses these viruses include the unassigned taxon HetrV6 and the new putative species of the genus *Mitovirus*. This observation is noteworthy as the entire reproductive cycle of mitoviruses is thought to occur inside host mitochondria (King et al., 2012; Hillman & Cai, 2013). *Thanatephorus cucumeris* mitovirus, which moderately resembles HetMV1-an1 (~24 % sequence identity at the protein level), is integrated into the host genome (Lakshman et al., 1998), but this appears not to be the case for HetMV1-an1. No members of the virus family *Narnaviridae* have been described from species of *Heterobasidion* despite extensive screening (Vainio et al., 2010, 2011a, b, 2012, 2013a, b, 2014; Kashif et al, 2014). Thus, the power of the small RNA deep sequencing approach in screening for new viruses is apparent.

**METHODS**

**Fungal strains and viral infections.** Isolates of *Heterobasidion* were used for the small RNA analysis: *H. abietinum* 04188 and RF252; *H. annosum* 94233 and 05021; *H. illgulare* 704; and *H. parvum* 7R18, 10084, RT3.49C, 95122 and 7R226. Seven of these isolates were known to be infected with partitiviruses or HetrV6 (Table 1), whereas three (05021, 10084 and RF252) contained no known virus infections. *H. parvum* 7R226 was known to be co-infected with strains of two viruses: Heterobasidion partitivirus 7 (strain HettPV7-pa1) and HetRV6 (strain HettRV6-pa19-c).

**Isolation and high-throughput sequencing of small RNAs.** Fungal mycelia were cultivated on modified orange agar plates covered with cellulose membranes. Hyphae were collected, frozen with liquid nitrogen and ground with a mortar and pestle. Total RNA was isolated from ~30-50 mg ground fungal mycelia using TRI Reagent (Molecular Research Center) and dissolved in sterile water that had been treated with diethylpyrocarbonate. The quantity and quality of the RNA samples were measured with agarose gel electrophoresis and a spectrophotometer (Qubit fluorometer; Invitrogen) and equal amounts of total RNA from ten *Heterobasidion* isolates were pooled. The quality of the pooled RNA sample (5 μg) was verified by the Fasteris SA Company, Switzerland, using a BioAnalyser 2100 (Agilent Technologies). The pooled sample had an rRNA (28S:18S) ratio of ≥2.6 and RNA integrity number of ≥6.9.

The pooled RNA sample was used for siRNA library construction by Fasteris SA, followed by high-throughput sequencing using an Illumina HiSeq 2500. Briefly, the sample processing included acrylamide gel purification of small RNA fragments, single-stranded ligation of 3’- and 5’-adapters, reverse transcription and PCR amplification to generate a cDNA library.

**Isolation of dsRNA with cellulose affinity chromatography.** The presence of mitoviral dsRNA in *H. annosum* 94233 was examined with CF11 cellulose affinity chromatography (Morris & Dodds 1979; Tuomivirta et al., 2002; Juvansuu et al., 2014) using 2.5-3.0 g (fresh mass) fungal mycelia. The protocol included isolation of total nucleic acids using phenol/chloroform extraction followed by specific precipitation of the dsRNA fraction using CF11 cellulose powder (Whatman) in 15 % (v/v) ethanol.

**Mitovirus sequence characterization with Sanger sequencing.** The complete genome of the novel putative mitovirus strain, HetMV1-an1, was determined with Sanger sequencing. cDNA was generated as described (Vainio et al., 2013a). Briefly, the protocol included isolation of total nucleic acids with phenol/chloroform extraction and precipitation with polyethylene glycol, followed by cDNA synthesis using random priming and the RevertAid reverse transcriptase (Thermo Scientific). PCR primers were designed based on siRNA contigs with mitoviral affinities as identified using Velvet (Zerbino & Birney, 2008; see below) and BLASTX (National Center for Biotechnology Information; www.ncbi.nlm.nih.gov). The following primers were used for sequence determinations: HetMV1MidF (5’-CTCTGTGAGGCACTTGGACC-3’), HetMV1MidR (5’-TTTTTGCAATTCAGGGCCGAT-3’), MV1MidDB (5’-GCCGGCACAAGCTATT-3’), MV1MidRecC (5’-CAAGTCTGTCATGTTTCA-3’), MV1MidRc (5’-GGAATTGGGTGGTGGTCTA-3’), MV1BeGev (5’-CTATCTCGCGCAAAAGCA-3’), MV1MidF (5’-ACCTGTGGCAGGCCACTGCGT-3’), MV1BeGevR (5’-ATGGCCCTGAATAGTGTGTT-3’), MV1BeR (5’-TATGCCATCGGAGGA-3’), MV1B (5’-AACCCTGAAAACGGGAAATCA-3’), MV1EF2 (5’-TCAAGGTGAAAAATGCCATGCA-3’), MV1BeR (5’-TATGGCCATCGGAACTGGA-3’).

**Quantification of viral dsRNAs and transcripts with reverse transcription-quantitative PCR (RT-qPCR).** For total RNA extraction, fungi were grown on cellulose-membrane-covered modified orange agar plates for 3–5 days. Total RNA was isolated from homogenized mycelia using TRI Reagent. cDNA was produced from 2 μg total RNA using the RevertAid First Strand cDNA Synthesis kit (Thermo Scientific) and random hexamer primers (Thermo Scientific).

dsRNA was isolated using cellulose affinity chromatography from fungi grown for 14–20 days on cellulose-membrane-covered modified orange serum agar plates as described by Juvansuu et al. (2014). Amplification of the adaptor oligonucleotide (5’-TTTGAT-CGGAGGCTGTTAAGC-3’) and the 5’-phosphorylated oligonucleotide (5’-GCTACTCCAGCCTGATACCA-3’) (100 μM each) was performed in 10 mM Tris (pH 8), 50 mM NaCl and 1 mM EDTA.
at 95 °C for 5 min, followed by slow cooling to room temperature. T4 DNA ligase (Promega) was used to ligate annealed dsDNA adaptors to the isolated dsRNA. Ligation products were purified using the E.Z.N.A. Cycle Pure kit (Omega bio-tek). The Maxima H Minus double-stranded cDNA synthesis kit (Thermo Scientific) was used for synthesis of double-stranded cDNA using the oligonucleotide 5′-TTGTATCGGCTTGTAAGG-3′. The reactions were treated with RNase I, cleaned with reagents from the E.Z.N.A. Cycle Pure kit and used for qPCR as described below.

The cDNAs were diluted with an equivalent amount of nuclease-free water before amplification with qPCR using EvaGreen enzyme (Solis BioDyne) in a Rotor-GeneQ machine (Qiagen). The following qPCR primers were used: GAPDH (Lehr et al., 2009), HetPV1-an1 RdRp forward (5′-GGTCAGCTCTGTGATCCGA-3′) and reverse (5′-AATTGTGGTGGCATGTAGCC-3′), HetPV1-an1 CP forward (5′-CAAACCA-ACCTGGCAGACTT-3′) and reverse (5′-GAGTCATCCACAGGAA-GA-3′), HetMV1-an1 RdRp forward (5′-ATTACCTCGGGTGTGGCCTC-3′) and reverse (5′-TGGTCCATGCTTTCCTTAAAT-3′), HetRV6-pa19-c RdRp forward (5′-ATGCAGCGCGACGTGCTGCC-3′) and reverse (5′-CGAGATCACGTGAGAAATGGAAGT-3′) and HetRV7-pa1 RdRp forward (5′-TATGTCGTCCTCCTGCCGCT-3′) and reverse (5′-GCAAGAGTTTACGAGGGAAGTT-3′). qPCR standards for each primer pair were obtained by cloning the respective PCR product into the vector TOPO-pCR2.1, using the TOPO TA Cloning kit (Invitrogen), and preparing plasmids containing preparations of 50 and 500,000 copies of the insert per 1 μl. All qPCR analyses were performed in triplicate as described by Jurvansuu et al. (2014). Bioinformatics. The small RNA contigs were assembled using the Velvet sequence assembler program (Zerbino & Birney, 2008) using k-mer values ranging from 13 to 25. No contigs were obtained using k-mer values smaller or larger than this. Host rRNA sequences (18S rRNA, ITS1, 5.8S rRNA, ITS2 and 28S rRNA) were removed from the resulting contigs. Velvet sequence assembler program (Zerbino & Birney, 2008) using k-mer values ranging from 13 to 25. No contigs were obtained using k-mer values smaller or larger than this. Host rRNA sequences (18S rRNA, ITS1, 5.8S rRNA, ITS2 and 28S rRNA) were removed from the resulting contigs. velvet sequence assembler program (Zerbino & Birney, 2008) using k-mer values ranging from 13 to 25. No contigs were obtained using k-mer values smaller or larger than this. Host rRNA sequences (18S rRNA, ITS1, 5.8S rRNA, ITS2 and 28S rRNA) were removed from the resulting contigs. velvet sequence assembler program (Zerbino & Birney, 2008) using k-mer values ranging from 13 to 25. No contigs were obtained using k-mer values smaller or larger than this. Host rRNA sequences (18S rRNA, ITS1, 5.8S rRNA, ITS2 and 28S rRNA) were removed from the resulting contigs. velvet sequence assembler program (Zerbino & Birney, 2008) using k-mer values ranging from 13 to 25. No contigs were obtained using k-mer values smaller or larger than this. Host rRNA sequences (18S rRNA, ITS1, 5.8S rRNA, ITS2 and 28S rRNA) were removed from the resulting contigs. velvet sequence assembler program (Zerbino & Birney, 2008) using k-mer values ranging from 13 to 25. No contigs were obtained using k-mer values smaller or larger than this. Host rRNA sequences (18S rRNA, ITS1, 5.8S rRNA, ITS2 and 28S rRNA) were removed from the resulting contigs.

Sequence-guided assembly of siRNAs was conducted using the MAQ program (Li et al., 2008). The percentage coverage of small RNAs mapped to reference genome segments (referred to as sequence coverage) was determined using alignments between the original reference sequences and the MAQ output sequences indicating nucleotide sites covered by siRNAs.

REFERENCES


Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using


