Open access • Posted Content • DOI:10.1101/2020.09.08.288829

# Discovery of divided RdRp sequences and a hitherto unknown genomic complexity in fungal viruses — Source link

Yuto Chiba, Takashi Yaguchi, Syun-ichi Urayama, Daisuke Hagiwara ...+1 more authors

**Institutions:** University of Tsukuba, Chiba University

**Published on:** 09 Sep 2020 - bioRxiv (Cold Spring Harbor Laboratory)

**Topics:** Fungal Viruses, RNA silencing, RNA, RNA polymerase and Gene

# Discovery of divided RdRp sequences and a hitherto unknown genomic complexity in fungal viruses

Yuto Chiba[1], Takashi Yaguchi[2], Syun-ichi Urayama[1,3,*], Daisuke Hagiwara[1,2,3,*]

[1] Laboratory of Fungal Interaction and Molecular Biology (donated by IFO), Department of Life and Environmental Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577, Japan

[2] Medical Mycology Research Center, Chiba University, 1-8-1, Inohana, Chuo-ku, Chiba 260-8673, Japan

[3] Microbiology Research Center for Sustainability (MiCS), University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8577, Japan

* Corresponding authors E-mail: urayama.shunichi.gn@u.tsukuba.ac.jp (SU); hagiwara.daisuke.gb@u.tsukuba.ac.jp (DH)

KEY WORDS: RNA virus, viral genome, FLDS, RdRp, Aspergillus fumigatus

## Abstract

By identifying variations in viral RNA genomes, cutting-edge metagenome technology has potential to reshape current concepts about the evolution of RNA viruses. This technology, however, cannot process low-homology genomic regions properly, leaving the true diversity of RNA viruses unappreciated. To overcome this technological limitation we applied an advanced method, Fragmented and Primer-Ligated Double-stranded (ds) RNA Sequencing (FLDS), to screen RNA viruses from 155 fungal isolates, which allowed us to obtain complete viral genomes in a homology-independent manner. We created a high-quality catalog of 19 RNA viruses (12 viral species) that infect *Aspergillus* isolates. Among them, nine viruses were not detectable by the conventional methodology involving agarose gel electrophoresis of dsRNA, a hallmark of RNA virus infections. Segmented genome structures were determined in 42% of the viruses. Some RNA viruses had novel genome architectures; one contained a dual methyltransferase domain and another had a separated RNA-dependent RNA polymerase (RdRp) gene. A virus from a different fungal taxon (*Pyricularia*) had an RdRp sequence that was separated on different segments, suggesting that a divided RdRp is widely present among fungal viruses, despite the belief that all RNA viruses encode RdRp as a single gene. These findings illustrate the previously hidden diversity and evolution of RNA viruses, and prompt reconsideration of the structural plasticity of RdRp. By highlighting the limitations of conventional surveillance methods for RNA viruses, we showcase the potential of FLDS technology to broaden current knowledge about these viruses.

42

43 **Author Summary**

44 The development of RNA-seq technology has facilitated the discovery of RNA
45 viruses in all types of biological samples. However, it is technically difficult to detect
46 highly novel viruses using RNA-seq. We successfully reconstructed the genomes of
47 multiple novel fungal RNA viruses by screening host fungi using a new technology,
48 FLDS. Surprisingly, we identified two viral species whose RNA-dependent RNA
49 polymerase (RdRp) proteins were separately encoded on different genome segments,
50 overturning the commonly accepted view of the positional unity of RdRp proteins in
51 viral genomes. This new perspective on divided RdRp proteins should hasten the
52 discovery of viruses with unique RdRp structures that have been overlooked, and
53 further advance current knowledge and understanding of the diversity and evolution of
54 RNA viruses.

55

## Introduction

Mycoviruses, one of the most extensively studied viral groups within the virosphere, infect fungi. These viruses inhabit the insides of the rigid fungal structure, and are transmitted to other fungal cells through cell division, sporogenesis, or cell-to-cell fusion (hyphal anastomosis), and their extracellular phase is hardly ever observed (Ghabrial and Suzuki, 2009). While there are no reports of mycoviruses killing their host fungi during the fungal life cycle, some asymptomatic infections appear to be reminiscent of symbiotic relationships. Surveillance studies have revealed that mycoviruses are, to a certain extent, present in isolates of plant pathogenic fungi (Arjona-Lopez et al., 2018; Urayama et al., 2010; van Diepeningen et al., 2006). In fact, more than 200 viral species have been detected in fungi to date (Gilbert et al., 2019).

Double-stranded (ds) RNA (dsRNA) has traditionally been used as the hallmark of RNA mycovirus detection. The rapid and specific extraction method for dsRNA from fungal cells and the conventional detection method, agarose gel electrophoresis (AGE) have accelerated large-scale mycovirus screening (Morris and Dodds, 1979; Okada et al., 2015) (Arjona-Lopez et al., 2018; Khankhum et al., 2017). However, AGE cannot detect low-level mycovirus infections because its sensitivity of viral detection is relatively low.

The establishment of metagenomic and metatranscriptomic analyses for virus surveillance over the last decade has enabled more viruses to be detected in ecologically diverse environments as well as in all living creatures. These methods, which rely on deep-sequencing analysis, were expected to be more sensitive than AGE at detecting

mycoviruses. Indeed, Illumina sequencing technology has successfully identified mycoviruses from fungal isolates previously considered free of these viruses by AGE (Nerva et al., 2016). Deep sequencing methods, however, are limited in their ability to identify novel viral sequences that lack homology to known virus-related sequences. Thus, finding a homology-independent method capable of detecting unidentified viruses and viral sequences in samples is required to fill this gap in viral sequence detection. Such a method would expand the list of virus-related sequences and augment current knowledge about viruses.

The multi-segmented RNA genomes possessed by some mycoviruses are coordinately replicated in the host. One segment contains an open reading frame (ORF) encoding an essential and universally seen RNA-dependent RNA polymerase (RdRp) in RNA viruses. Additional ORFs are contained in the other segment(s). In most cases, the cognate segments can be identified via the conserved sequences at their terminal ends (Urayama et al., 2018), but the conventional high-throughput sequencing methods mostly suffer from a loss of information for the terminal sequences. This technical limitation has hindered the discovery of the cognate segments and the novel viral sequences on them. To overcome this issue, we have recently developed ‘Fragmented and primer-Ligated DsRNA Sequencing (FLDS)’ technology to enable researchers to identify complete viral RNA genomes (Urayama et al., 2018; Urayama et al., 2016). This method provides reliable terminal sequences for each genome, from which the segmented RNA genomes of viruses can be determined in a homology-independent manner. This technology has higher detection power than other current methods. These

advantages have allowed us to obtain highly informative complete genomes and identify uncharacterized mycoviruses using FLDS.

In the present study, we adopted FLDS technology to comprehensively screen for mycoviruses in *Aspergillus* species. Altogether, 155 isolates of *A. fumigatus* and its related species were screened, and the complete genomes of 18 viruses, as well as one with an incomplete genome, were determined for 16 fungal isolates. The FLDS-based high quality sequences we obtained supported the multi-segmented genome structure for eight viruses and uncovered novel viral sequences containing seven predicted novel ORFs in total. The most surprising finding from this screening is that one virus carries a partial RdRp lacking the essential C/D motif. The cognate segment of this virus encodes a novel ORF containing a C/D-like sequence motif. These unexpected results confirm that the ability to obtain complete, high quality viral genomes makes FLDS technology a powerful tool for expanding our current understanding of diversity in RNA viruses.

## Results

**Comprehensive screening for RNA viruses using FLDS.** Altogether, 155 *Aspergillus* strains were used to screen for RNA viruses (Table S1). First, the dsRNA extracted from the mycelium of each strain was subjected to AGE. As a result, dsRNA bands were visibly detected from four *A. fumigatus* and five *A. lentulus* strains (Fig. 1). The patterns of the dsRNA bands differed from each other. The dsRNAs from the remaining 146 strains were pooled into eight groups (dsRNAs from ~20 strains per group) and

sequenced using FLDS to identify viruses that were undetectable by AGE. After RT-PCR verification, viral sequences were identified in eight strains. Overall, 17 fungi were identified that were infected with an RNA virus (Table 1), and the frequency of RNA virus-positive isolates of each species was as follows: *A. fumigatus* 8/79 (10.1%), *A. lentulus* 5/27 (18.5%) and *A. pseudoviridinutans* 4/15 (26.7%), *A. udagawae* 0/19 (0%), and *Neosartorya fischeri* 0/15 (0%) (Fig. S1). The *A. fumigatus* set, which included 20 environmental and 59 clinical isolates, contained three and five isolates that were infected with RNA viruses, respectively.

**Determination of segmented virus genomes.** We determined the full-length viral sequences by FLDS for the strains that we identified by AGE and the pooled FLDS. Altogether, we determined the complete sequences of 16 fungal strains within three species, *A. fumigatus*, *A. lentulus*, and *A. pseudoviridinutans*. The complete sequence from *A. pseudoviridinutans* IFM 59503 could not be obtained despite this virus being detected by RT-PCR. The segmented viral genomes were determined according to the terminal sequence similarity among them (Fig. S2). For example, an RNA virus isolated from *A. fumigatus* IFM 62632 was found to contain four segments, whereas *A. fumigatus* IFM 63431 and IFM 62629 each contained three segments (Table 1). As a consequence, clarification of each segment resulted in us being able to differentiate the co-infecting viruses in the fungal isolates. In fact, *A. fumigatus* IFM 62632 and IFM 63147 were co-infected with two and three viruses, respectively. Altogether, 20 viruses were identified in 17 *Aspergillus* isolates, among which eight had segmented genomes

144 (two each for bi-segments, four each for tri-segments, and two each for tetra-segments)
145 (Table 1). The FLDS method detected the viral genomes with high sensitivity, even
146 when multiple viruses had co-infected the same host, and it accurately discriminated
147 segmented genomes in the viruses.
148
149 **RNA virus classification.** Our BlastX analysis revealed that all 20 of the viral RNA
150 sequences contained RdRp (Table S4). Based on the criterion that the RdRp-encoding
151 segments sharing > 90% nucleic acid sequence identity were recognized as single
152 operational taxonomic units (OTUs), 20 viruses fell into 12 OTUs. The virus from *A.*
153 *pseudoviridinutans* IFM 59503 with a partial sequence also fell into one of the OTUs.
154 Based on the high sequence identity (> 95%) shared with known viral sequences, four
155 viruses were recognized as Aspergillus fumigatus polymycovirus 1 (AfuPmV1),
156 Aspergillus fumigatus chrysovirus (AfuCV), Aspergillus fumigatus narnavirus 2
157 (AfuNV2), and Aspergillus fumigatus mitovirus 1 (AfuMV1). The other eight OTUs
158 were regarded as novel viral species, and were therefore tentatively named according to
159 the taxonomical linage of the top 'hits' for RNA viruses in the BlastX analysis against
160 the NCBI nr database (Table S4). Six of eight viral species were related to dsRNA virus
161 families (*Partitiviridae* and *Totiviridae*), positive ssRNA virus families (*Narnaviridae*
162 and *Botourmiaviridae*), and a negative ssRNA virus family (*Betamycobunyaviridae*, a
163 previously suggested virus family), whereas two OTUs were considered to be
164 unclassified RNA viral linages. Notably, six of the 20 viruses identified herein, all of
165 which had been detected by AGE, are considered to be dsRNA viruses.

166 *Aspergillus fumigatus* is a relatively well-studied fungal species in terms of
167 mycovirus screening; nonetheless, three viruses, Aspergillus fumigatus botourmiavirus
168 1 (AfuBOV1), Aspergillus fumigatus negative-strand RNA virus 1 (AfuNSRV1), and
169 Aspergillus fumigatus RNA virus 1 (AfuRV1), are newly identified by our FLDS-based
170 screening. Viral isolation from related fungal species, such as *A. lentulus* and *A.*
171 *pseudoviridinutans*, has never been reported; hence, the following viruses identified
172 from the fungal species are new species: Aspergillus lentulus partitivirus 1 (AlePV1),
173 Aspergillus lentulus non-segmented dsRNA virus 1 (AleNdsRV1), Aspergillus lentulus
174 narnavirus 1 (AleNV1), Aspergillus lentulus totivirus 1 (AleTV1), and Aspergillus
175 pseudoviridinutans botourmiavirus 1 (ApvBOV1).
176
177 **Structures of the novel RNA viruses identified in *Aspergillus* fungi.** AfuRV1, which
178 was identified in *A. fumigatus* IFM 63439, consists of three RNA segments (3,611,
179 3,447 and 1,943 nucleotides long, excluding the polyA-like region) (Fig. 2A). Each
180 segment contains a single ORF lacking significant nucleotide identity to known
181 sequences in the NCBI nt database. ORF1 contains methyltransferase (E-value =
182 $1.8\times10^{-10}$) and RdRp (E-value = $2.6\times10^{-82}$) domains, and ORF2 contains
183 methyltransferase (E-value = $3.3\times10^{-25}$) and helicase (E-value = $2.5\times10^{-21}$) domains.
184 BlastX analysis showed that the sequence with the top hit for ORF1 was RdRp from
185 Luckshill virus (LuV) (an unclassified ssRNA virus) (coverage, 93.0%; E-value, 0;
186 identity, 32.7%), whereas that for ORF2 was a hypothetical protein from Cyril virus
187 (another unclassified ssRNA virus) (coverage, 78.0%; E-value, $3.0\times10^{-40}$; identity,

188 32.2%). Phylogenetic analysis of the RdRp domain showed that AfuRV1 falls into the
189 virga-like virus clade of viruses previously isolated from invertebrates and fungi (Fig.
190 2B). AfuRV1, however, fell into the invertebrate-derived sub-clade containing LuV,
191 rather than the mycovirus sub-clade. Interestingly, viruses within the virga-like clade
192 have never been reported to possess segmented genomes. We found that the AfuRV1
193 genome contains an additional segment. Comparing the methyl transferase domains of
194 ORF1 and ORF2, ORF2 fell into the virga-like virus clade, but ORF1 did not (Fig. 2C).
195 The helicase domain in ORF2 did not fall within an established family or proposed
196 group (Fig. 2D).

197 AleNV1, which originated from *A. lentulus* IFM 63547, contains two
198 segmented genomes (3,071 and 1,814 nucleotides long, excluding the polyA-like
199 region) (Fig. S3), whose RdRp genes share low sequence homology with that of the
200 Beihai narna-like virus 21. The ORFs in RNA2 did not share significant similarity with
201 known proteins. Based on its RdRp sequence, AleNV1 belongs to the well-established
202 *Narnavirus* genus (Fig. 3). No viruses with bi-segmented genomes have been reported
203 so far in viral genera, except LepseyNLV1 (whose complete sequence has not been
204 reported), and MaRNAV1, which was isolated from the human trypanosomatid parasite
205 *Leptomonas seymouri* (Grybchuk et al., 2018) and the human malaria parasite
206 *Plasmodium vivax* (Charon et al., 2019). Thus, AleNV1 has a unique genome structure
207 and is the first reported isolate with a bi-segmented genome among mycoviruses from
208 the *Narnavirus* genus.
209

**Discovery of a novel divided RdRp sequence.** AfuNV2, a previously reported virus (Zoll et al., 2018), was identified in three different *A. fumigatus* isolates (IFM 63147, IFM 63431 and IFM 62629) (Table 1). Although the sequence reported by another research group is non-segmented, our FLDS-based sequencing revealed that AfuNV2 possesses RNA2 and RNA3 segments in addition to RNA1 (RdRp). This tri-segmented genome was confirmed to contain the highly conserved terminal sequences described above (Fig. S2), and was also confirmed by AGE where the band patterns were seen to correspond to the tri-segmented genome's length (Fig. 1 lane 2 and 3). RNA2 and RNA3 are predicted to encode a large single ORF (ORF2) and multiple short ORFs (ORF3, ORF4, and ORF5) (Fig. 4A). BlastP analysis showed that the sequence with the top hit for ORF2 was RdRp from *Plasmopara viticola*-associated narnavirus 33 (PvaNV33) (coverage, 74.0%; E-value, 3E-17; identity, 23.9%), whereas ORF3, ORF4, and ORF5 share no significant similarities (e-value $\leq 1\times10^{-5}$) with known protein sequences in the NCBI nr database or in the Pfam domain database.

When the RdRp amino acid sequences were aligned, we noticed that the RdRp-encoding ORF1 from AfuNV2 lacked motifs C and D, but the well-conserved motifs (F, A and B) among the narnaviruses from other fungi were present (Fig. 4B, 4C). Interestingly, closely related narnaviruses (CtNV1, AtNV1, NpNV2, FuPNV2 and AfuNV1) were also found to lack the C and D motifs (Fig. 4C) (Lin et al., 2020). In motif C, GDD, a conserved amino acid sequence, is believed to play an essential catalytic role. Therefore, we searched for the GDD sequence in the other RNA2 and RNA3 ORFs. After much searching, we found a potential sequence in the N-terminal

232 region of the RNA2 ORF. Sequence alignments suggested that this region includes
233 amino acid residues that are conserved in the C and D motif regions of the RdRp
234 proteins from narnaviruses (Fig. 4C). This was observed in three of the sequences from
235 the AfuNV2 viruses we isolated. However, we were unable to confirm whether the
236 aforementioned closely related CtNV1, AtNV1, NpNV2, FuPNV2 and AfuNV1 species
237 contain RNA2 with an ORF containing a C/D-like motif sequence, because their
238 sequences were deposited as non-segmented genomes.

239 To support this finding, we searched for other viruses that lack the C and D
240 motifs in RdRp. We eventually found one known RNA virus by FLDS sequencing,
241 Magnaporthe oryzae narnavirus 1 (MoNV1) (MN480844) from *Pyricularia*
242 (*Magnaporthe*) *oryzae* APU10-199A, with a complete genome and reliable terminal
243 sequences in each segment. While a non-segmented genome was reported for MoNV1,
244 the MoNV1 we isolated has four genomic segments, but motifs C and D are missing
245 from RdRp on RNA2 (Fig. 4D). The C/D-like motif lies within the N-terminal region of
246 the ORF of RNA1, as was the case of AfuNV2 RNA2 (Fig. 4B). The ORFs on RNA3
247 and RNA4 from MoNV1 contain no predicted proteins or domains with significant
248 homology. In addition to MoNV1, another narnavirus, MoNV2, co-infected the same
249 strain. MoNV2 has a non-segmented genome containing a single ORF that shares
250 significant nucleotide identity with RdRp from AleNV1 (coverage, 99.0%; E-value, 0;
251 identity, 74.61%) (Fig. 3). The RdRp from MoNV2 harbors a C/D motif (Fig. 4C).
252 Collectively, the results for these two different fungal species (*A. fumigatus* and *P.*
253 *oryzae*) suggest that certain narnaviruses carry RdRp sequences that are divided into

254 two different ORFs, with one ORF containing motifs F, A and B, and the other
255 potentially containing motifs C and D.
256
257 **Discussion**
258 DsRNA mycoviruses infect a wide range of fungal species including
259 Ascomycota, Basidiomycota, Glomeromycota, and Mucoromycotina (Gilbert et al.,
260 2019, Turina et al. 2018, Kartali et al. 2019). Some viruses can affect growth,
261 development, toxin production, and pathogenicity in their fungal hosts. Fungal viruses
262 have not been intensively studied despite their potential impact on ecology, agriculture,
263 food security, and public health. In this study, we evaluated the recently developed
264 FLDS technology for its sensitivity and accuracy of RNA viral sequence detection.
265 Several new viral sequences were successfully identified as complete genomes. A most
266 intriguing finding was the presence of a divided RdRp gene. No examples like this
267 irregular form have been reported in any other RNA viruses known to infect living
268 creatures. This result could be only accomplished by FLDS-based sequencing because
269 of its ability to produce high-quality end-to-end sequences. This demonstrates its utility
270 for capturing previously unidentified viral sequences from RNA viruses.
271 Large-scale AGE screening for RNA viruses was conducted on a set of more
272 than 300 *Aspergillus* strains (Bhatti et al., 2012), and 6.6% of the detected strains
273 contained dsRNA. Elsewhere, of the 86 *A. fumigatus* clinical isolates that were
274 examined 18.6% contained dsRNA (Refos et al., 2013). Here, we detected dsRNA by
275 AGE from 9/155 strains (5.8%) of *Aspergillus* species. Another advanced screening

method, metatranscriptomics, was recently used to detect viral sequences, and 72 assembled virus-related sequences from 275 isolates of plant pathogenic fungi were identified (Marzano et al., 2016). The searches were based on identifying sequences with significant identity to known amino acid-encoding viral sequences. Notably, only 15% of the sequences were predicted to be derived from dsRNA, while 73% and 12% were predicted to be derived from positive-sense RNA or negative-sense RNA viruses, respectively. More recently, a pipeline to efficiently detect viral sequences from a transcriptomics dataset was proposed (Gilbert et al., 2019). This was also based on sequence similarity to RdRp. Of the 569 RNA-Seq samples, 59 complete mycoviral genomes were identified in 47 datasets, 34 viruses (57%) were predicted to be dsRNA viruses, and 88% were new species. Here, 10 viruses were identified by FLDS that were overlooked by AGE. All the overlooked viruses were predicted to be ssRNA viruses. We consider that the amount of dsRNA recovered as a replicative intermediate of ssRNA is low. Nonetheless, FLDS captured sequences derived from ssRNA viruses. Therefore, the highly sensitive, high-throughput sequence-based screening of FLDS is a powerful tool for constructing deep and wide viral catalogs.

One apparent limitation occurs with homology-based viral sequence detection whereby the novel ORFs accumulated by knowledge-based updating can overlook those lacking homology to known virus-related ORFs, resulting in a biased list of viral sequences. Our FLDS-based screening was clearly able to circumvent this issue by identifying six viral species with segmented genomes and discovering seven ORFs that had never been recognized as virus-related sequences (ORFs encoding hypothetical

298 proteins with unknown functions, Table S4).

299 The discovery of segmented genomes provides evolutionary insight into the
300 origin of certain mycovirus groups. AfuRV1 has three segmented genomes and contains
301 two methyltransferase domains in different segments. As far as we are aware, RNA
302 viruses with multiple methyltransferase domains have never been reported in published
303 literature. Phylogenetic analysis revealed that the methyltransferase domain of ORF2
304 (MT2) and the RdRp domain of ORF1 belong to a virga-like virus clade. In contrast, the
305 methyltransferase domain of ORF1 (MT1) and the helicase domain of ORF2 fall into an
306 unclassified group (not the virga-like virus clade). This suggests that the ancestor of
307 AfuRV1 acquired MT1 and its helicase domain from a different virus. Identifying the
308 origins of these domains and the evolutional history of AfuRV1 is not straightforward
309 because MT1 and the helicase domain are distantly related to the domains of known
310 viruses. By FLDS analysis, we identified segmented genomes in AfuNV2 and AleNV1
311 that both belong to the *Narnaviridae* family. Interestingly, segmented genomes have
312 only been reported for LepseyNLV1 (a protozoal virus) and *Botourmiaviridae* family
313 plant viruses, and they have not been reported among the *Narnaviridae* family or
314 closely-related ssRNA viruses (Grybchuk et al., 2018; Rastgou et al., 2009). Hence, this
315 is the first identification of *Narnaviridae* family mycoviruses with multi-segmented
316 genomes. It is noteworthy that AfuNV2 and AleNV1 were not classified as belonging to
317 the sub-clade that includes LepseyNLV1. Furthermore, AfuNV2 and AleNV1
318 narnaviruses are not closely related to each other and have different genome structures.
319 This suggests that during evolution of the *Narnaviridae* family and its relative ssRNA

320 viruses, the acquisition of segments and changes in genome structure occurred
321 independently in each host kingdom. We cannot, however, rule out the possibility that
322 non-RdRp encoding segments are satellite RNAs that tentatively coexist. Further
323 investigations in this area are required.

324 The gene encoding RdRp is universally present among RNA viruses (Wolf et
325 al., 2018), and all known RNA viruses encode RdRp in a single ORF (King et al., 2012).
326 Unexpectedly, we found that viruses in a certain clade of *Narnaviridae* encode an RdRp
327 that lacks catalytic domains C and D, and a different coexisting ORF encodes the
328 missing domains. This irregular genome structure was identified in two different viruses
329 from two different fungal hosts, *Aspergillus* and *Pyricularia*. Besides these two viruses,
330 this group includes other viruses isolated from *Fusarium* (FuPNV2), *Neofusicoccum*
331 (NpNV2), *Cladosporium* (CtNV1), and *Alternaria* (AtNV1) (Fig. 3). According to the
332 deposited sequences, these viruses appear to lack domains C and D of RdRp. These
333 sequences were not obtained by FLDS; thus the corresponding ORF with a GDD motif
334 or cognate genomes were unlikely to be detected; indeed, no such sequences have been
335 deposited. The deposited sequence dataset from other narnaviruses supports the
336 possibility that imperfect RdRp proteins exist in a certain group of mycoviruses that
337 infect a wide range of the Ascomycetes phylum. As described above, the ORF2 of
338 AfuNV2 showed low but certain similarity to the RdRp sequence from PvaNV33.
339 Interestingly, PvaNV33 showed certain similarities not only to ORF2 but also to ORF1
340 of AfuNV2 (Fig. S4). Moreover, sequences of RdRp from PvaNV32 (GenBank:
341 QIR30311.1) and PvaNV35 (GenBank: QIR30314.1) that formed a single clade with

PvaNV33 also shared certain similarities to both ORFs of AfuNV2. This complementary genomic structure suggests that the RdRps from these *Plasmopara viticola*-associated narnaviruses are the molecular ancestors from which ORF1 and ORF2 of AfuNV2 were derived by division. During the preparation of this manuscript, another research group reported on the identification of Magnaporthe oryzae narnavirus virus 1 (MoNV1) from an *M. oryzae* strain isolated in China (Lin et al. 2020). The viral sequence was deposited as a single genome with RdRp lacking domains C and D, which concords with our results for the MoNV1 we identified in our work. This consistency in the viral sequences from different countries suggests that viruses with irregular genome structures are widely distributed, meaning that the event leading to the atypical structures did not occur locally. Our finding of divided RdRp sequences raises questions about how divided RdRps function in host cells and whether atypical structures are limited to mycoviruses. Answering these questions is of great interest to us.

In conclusion, our FLDS-based screening for mycoviruses using 155 fungal isolates has led to the discovery of novel species and novel segmented genomes in some of the viruses. Some viral genomes have novel structures that would not have been captured by conventional methods. Thus, FLDS-based screening has potential to tap into unexplored diversity in RNA viruses.

## Materials & Methods

### Strains and culture conditions

The *Aspergillus* strains used in this study are listed in Table S1. These strains were

364 cultured in potato dextrose broth (PDB) with reciprocal shaking (120 rpm) for up to 5
365 days at 30°C or 37°C. All the strains were provided by the National BioResource
366 Project (https://nbrp.jp/). *Pyricularia* (*Magnaporthe*) *oryzae* APU10-199A (Higashiura
367 et al., 2019) was cultured in PDB for 1 week at 30°C before harvesting.
368
369 RNA extraction
370 Fungal mats (fresh weight, 100 mg) were disrupted in liquid nitrogen in a mortar or
371 using FastPrep 24 (MP Biomedicals Inc., OH, USA). dsRNA and ssRNA purification
372 was performed as described previously (Urayama et al., 2020; Urayama et al., 2018;
373 Urayama et al., 2016). In brief, total nucleic acids were manually extracted from the
374 ground cells with sodium dodecyl sulfate–phenol. dsRNA was purified using the
375 cellulose resin chromatography method (Morris and Dodds, 1979; Okada et al., 2015)
376 and subjected to AGE analysis. To obtain sequence-grade dsRNA, the remaining DNA
377 and ssRNA were removed with amplification grade DNase I (Invitrogen, Carlsbad, CA,
378 USA) and S1 nuclease (Invitrogen). Total RNA was extracted from the pulverized
379 samples using the TRIzol Plus RNA Purification Kit (Invitrogen), and the eluted RNA
380 was treated with amplification grade DNase I (Invitrogen) and purified using RNA
381 Clean & Concentrator-5 (Zymo research, Irvine, CA, USA).
382
383 Sample pooling, sequence library construction and sequencing
384 When dsRNA band(s) were visible by AGE, the dsRNA samples from each
385 isolate were individually prepared for viral genome sequencing by FLDS. When no

386 visible dsRNA band or bands were observed, the dsRNA samples from up to 20 isolates
387 were pooled into a single sample (referred to as pool 1–8), following FLDS analysis
388 (referred to as pooled-FLDS analysis).

389 dsRNA was converted into dscDNA by the FLDS method (Urayama et al., 2018).
390 Each purified dsRNA was fragmented by ultrasound using a Covaris S220
391 ultrasonicator (Woburn, MA, USA), and a U2 adapter was ligated to each dsRNA
392 fragment using T4 RNA ligase (Takara Bio Inc., Kusatsu, Japan). After denaturing the
393 product, single-stranded (ss) cDNA (sscDNA) was synthesized using the SMARTer
394 RACE 5′/3′ Kit (Takara) with a U2-complementary primer. dscDNA was obtained by
395 PCR with a U2-complementary primer and a universal primer mix (provided by the
396 SMARTer RACE 5′/3′ Kit).

397 cDNA libraries were constructed as described previously (Urayama et al., 2018).
398 Each 300 bp of the paired-end sequences from each fragment were determined on the
399 Illumina MiSeq platform (Illumina, CA, USA).

400

401 Data processing

402 Clean reads were obtained by removing low-quality, adapter and low-complexity
403 sequences as described previously (Urayama et al., 2018). For the RNA virome analyses,
404 contaminated rRNA reads were removed by SortMeRNA (Kopylova et al., 2012).
405 According to a previous method (Urayama et al., 2018), the cleaned reads were
406 subjected to *de novo* assembly using the CLC Genomics Workbench version 11.0 (CLC
407 Bio, Aarhus, Denmark). The resulting assemblies were manually examined and

408 extended using the assembly Tablet viewer (Milne et al., 2010). Where the terminal end
409 of a contig ended with same bases for more than 10 reads or a polyA sequence was
410 present, the position was recognized as the terminal end of the RNA genome. When a
411 contig had termini at both of its ends, it was considered to be the full-length sequence of
412 the RNA genome. Multi-segment genomes were judged according to the terminal
413 sequence similarities of the segments. BlastN and BlastX programs (Camacho et al.,
414 2009) were used to identify sequence similarities among known nucleotide sequences
415 and protein sequences, respectively.
416
417 Phylogenetic analysis
418 ORF prediction was performed by ORFfinder, after which Pfam domain
419 searching was conducted (Finn et al., 2016). Phylogenetic analysis of the helicase and
420 methyltransferase domains of RdRp, which were based on the amino acid sequences
421 obtained from the NCBI nr database, were aligned using MUSCLE (Edgar, 2004) in
422 MEGA6 (Tamura et al., 2013). The accession numbers of the sequences used for the
423 analyses are listed in Table S2. Alignment-ambiguous positions were removed with
424 trimAl (Capella-Gutiérrez et al., 2009). Maximum likelihood-based phylogenetic
425 analyses were performed using RAxML (Stamatakis, 2014), and bootstrap tests were
426 conducted with 1,000 samplings. The amino acid substitution model was selected by
427 Aminosan (Tanabe, 2011) using Akaike's information criterion (Sugiura, 1978). To
428 visualize phylogenetic trees, Fig-Tree (Rambaut, 2014) was used.
429

Reverse transcription (RT) PCR (RT-PCR)

To detect RNA viruses from host fungi in pooled samples, RT-PCR analyses were performed using specific primer sets (Table S3) as described previously (Urayama et al., 2014). After pricking the mycelia grown on a potato dextrose agar several times with a toothpick, the toothpick was then dipped into the one-step RT-PCR reaction mix in a PCR tube and twisted three times. One-step RT-PCR was performed using the Super-Script III One-Step RT-PCR System with Platinum Taq (Invitrogen) according to the manufacturer's protocol. To support the sequences provided by FLDS, the 3′ end sequences of RNAs 1 and 2 from Aspergillus fumigatus RNA virus 1 (AfuRV1) were confirmed by One-step RT-PCR using an oligo (dt) primer and specific primers (Table S3). Total RNA was used as the template. PCR products were run on 1% agarose gels and the visualized fragments were excised, purified using the FastGene Gel/PCR Extraction Kit (Nippon Genetics, Tokyo, Japan), and then used for direct Sanger sequencing.

Data accessibility

Data sets supporting the results of this study are available in the GenBank database repository (Accession Nos. DDBJ: LC553675–LC553714) and the Short Read Archive database (Accession No. DDBJ: DRA010415).

**Acknowledgments**

This research was supported by a grant from the Institute for Fermentation, Osaka, and in part by a Grant-in-Aid for Scientific Research (18H05368 and 20H05579) from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan and a Grant-in-Aid for Scientific Research on Innovative Areas from the Ministry of Education, Culture, Science, Sports, and Technology (MEXT) of Japan (No. 16H06429, 16K21723, and 16H06437). We thank Drs. Hiromitsu Moriyama and Shin-ichi Fuji for providing *M. oryzae* APU10-199A. We thank Sandra Cheesman, PhD, from Edanz Group (https://en-author-services.edanzgroup.com/ac) for editing a draft of this manuscript.

## Conflict of interest

The authors declare that there are no conflicts of interest.

## Figure Legends

**Fig. 1: Detection of dsRNA by agarose gel electrophoresis**. Comprehensive screening of 155 *Aspergillus* strains produced nine strains with positive virus-like bands. The dsRNAs from nine strains (four from *A. fumigatus* and five of *A. lentulus*) were electrophoresed, and the gel was stained with GelRed. Lanes: M, DNA marker; 1, IFM 62632; 2, IFM 63147; 3, IFM 63431; 4, IFM 64916; 5, IFM 62627; 6, IFM 64004; 7, IFM 64003; 8, IFM 63547; and 9, IFM 65052.

474 **Fig. 2: Characterization of AfuRV1.** (A) The RNA genome structure model for
475 AfuRV1. The predicted ORFs are indicated by white boxes. The domains identified as
476 methyl transferase, RdRP, and helicase are indicated in yellow, blue and brown boxes,
477 respectively. Molecular phylogenetic analysis on the RdRp (B), methyltransferase (C),
478 and helicase (D) domains was performed by maximum likelihood-based methodology.
479 The numbers indicate the percentage bootstrap support from 1,000 RAxML bootstrap
480 replicates. The best-fitting amino acid substitution models were [LG+F+G] (B) (C) and
481 [rtREV+F+G] (D). The accession numbers and full virus names are listed in Table S2.
482 The scale bar represents the number of substitutions per site. The viral sequences from
483 fungi and invertebrates are indicated by orange circles or blue squares, respectively.
484 AfuRV1 sequences are shown in red font.
485
486 **Fig. 3: Phylogenetic tree for the RdRp from the *Narnaviridae* family.** The RdRp
487 amino acid sequences without motifs C and D from AleNV1, AfuNV2, MoNV1, and
488 MoNV2 (without the C/D motif is shown in a gray box) and related viruses were used
489 to construct the phylogenetic tree. The numbers indicate the percentage bootstrap
490 support from 1,000 RAxML bootstrap replicates. The best-fitting amino acid
491 substitution model was [LG+F+G]. The accession numbers and full virus names are
492 listed in Table S2. The viral sequences from fungi, invertebrates, oomycetes, plants, and
493 an unknown host are indicated by orange circles, blue squares, light blue diamonds,
494 green triangles, and question marks, respectively. The viral sequences identified in this
495 study are shown in red font.

**Fig. 4: Characterization of AfuNV2 and MoNV1.** (A) RNA genome structure model for AfuNV2. The predicted ORFs are represented by boxes, and the first identified RdRp domain is shown by a blue box. (B) Schematic model of RdRp protein with conserved motifs in AfuNV2, MoNV1 and ScNV-20S genomes. Conserved A–D and F motifs in RdRp from ScNV-20S, a type strain of narnavirus, are shown (Venkataraman et al., 2018, Jia, H & Gong, P, 2019). (C) Multiple alignment of the deduced amino acid sequences of the RdRp motifs for the *Narnavirus* genus. Among the 18 to 24 sequences, the amino acid positions with 100% matches and >50% matches are depicted by green and gray shading, respectively. Dominant amino acid residues at positions with >50% amino acid matches are shown in bold. The number of amino acids not shown in the alignment is noted in each sequence. The accession numbers and full virus names are listed in Table S2. (D) RNA genome structure model for MoNV1. The predicted ORFs are represented by boxes, and the first identified RdRp domain is shown by a blue box.

**Fig. S1: Frequency of RNA virus-positive isolates and the detection method.** AGE: agarose gel electrophoresis.

**Fig. S2: Multiple alignments of the 5′- and 3′-terminal regions of the RNA sequences.** Nucleotide positions with 100% matches among the sequences are depicted by green shading. Numbers at the beginning and end of each sequence represent the

518 nucleic acid positions.

519

520 **Fig. S3: Genome structure models for the viruses identified in this study.** The
521 predicted ORFs are shown by a box. The domains identified in RdRp and the coat
522 protein (CP) are shown by blue and gray boxes, respectively. Hypothetical proteins are
523 shown as 'hypothetical'.

524

525 **Fig. S4: Schematic model of RdRp protein from PvaNV33 and comparison with**
526 **AfuNV2.** Conserved A–D and F motifs in RdRp from PvaNV33 and AfuNV2 are
527 shown. The regions with similarity between two proteins are shown in a striped square.

528

529 **Table 1:** List of *Aspergillus* strains carrying viral sequences

| Host fungal species and strain | Virus name | Number of segments | Genome type [*4] | Detection method [*5] |
|---|---|---|---|---|
| ***A. fumigatus* IFM 62632** | Aspergillus fumigatus polymycovirus 1 (AfuPmV1) | 4 | dsRNA | AGE |
| | Aspergillus fumigatus negative-stranded RNA virus 1 (AfuNSRV1) [*1] | 1 [*2] | ssRNA(-) | FLDS |
| ***A. fumigatus* IFM 63147** | Aspergillus fumigatus chrysovirus (AfuCV) | 4 | dsRNA | AGE |
| | Aspergillus fumigatus narnavirus 2 (AfuNV2) | 3 | ssRNA(+) | AGE |
| | Aspergillus fumigatus botourmiavirus 1 (AfuBOV1) [*1] | 1 | ssRNA(+) | FLDS |
| ***A. fumigatus* IFM 63431** | Aspergillus fumigatus narnavirus 2 (AfuNV2) | 3 | ssRNA(+) | AGE |
| ***A. fumigatus* IFM 64916** | Aspergillus fumigatus botourmiavirus 1 (AfuBOV1) [*1] | 1 | ssRNA(+) | AGE |
| ***A. fumigatus* IFM 62355** | Aspergillus fumigatus mitovirus 1 (AfuMV1) | 1 | ssRNA(+) | Pooled FLDS |
| ***A. fumigatus* IFM 62629** | Aspergillus fumigatus narnavirus 2 (AfuNV2) | 3 | ssRNA(+) | Pooled FLDS |
| ***A. fumigatus* IFM 63439** | Aspergillus fumigatus RNA virus 1 (AfuRV1) [*1] | 3 | ssRNA(+) | Pooled FLDS |
| ***A. fumigatus* IFM 64779** | Aspergillus fumigatus botourmiavirus 1 (AfuBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |
| ***A. lentulus* IFM 62627** | Aspergillus lentulus partitivirus 1 (AlePV1) [*1] | 2 | dsRNA | AGE |
| ***A. lentulus* IFM 63547** | Aspergillus lentulus narnavirus 1 (AleNV1) [*1] | 2 | ssRNA(+) | AGE |
| ***A. lentulus* IFM 64003** | Aspergillus lentulus non-segmented dsRNA virus 1 (AleNdsRV1) [*1] | 1 | dsRNA | AGE |
| ***A. lentulus* IFM 64004** | Aspergillus lentulus non-segmented dsRNA virus 1 (AleNdsRV1) [*1] | 1 | dsRNA | AGE |
| ***A. lentulus* IFM 65052** | Aspergillus lentulus totivirus 1 (AleTV1) [*1] | 1 | dsRNA | AGE |
| ***A. pseudoviridinutans* IFM 59502** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |
| ***A. pseudoviridinutans* IFM 59503** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1 [*3] | ssRNA(+) | Pooled FLDS |
| ***A. pseudoviridinutans* IFM 61377** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |
| ***A. pseudoviridinutans* IFM 61378** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |

530 *1 : Tentative virus names were used to represent the novel sequences identified in the strains from this study.
531 *2 : The terminal sequence was not completely determined.
532 *3 : The presence of the viral sequence was confirmed by RT-PCR. However, the complete sequence was not determined.
533 *4 : The genome type of each novel virus was predicted based on that of similar viruses.
534 *5 : AGE, agarose gel electrophoresis; FLDS: the dsRNA from the single isolate was analyzed by FLDS; Pooled FLDS: the dsRNA
535 samples from multiple (~20) isolates were pooled and analyzed by FLDS
536
537

538

539

540

541

542 **References**

543 Arjona-Lopez, J.M., Telengech, P., Jamal, A., Hisano, S., Kondo, H., Yelin, M.D.,
544 Arjona-Girona, I., Kanematsu, S., Lopez-Herrera, C.J., Suzuki, N., 2018. Novel,
545 diverse RNA viruses from Mediterranean isolates of the phytopathogenic fungus,
546 Rosellinia necatrix: insights into evolutionary biology of fungal viruses. Environ.
547 Microbiol. 20, 1464-1483.

548 Bhatti, M.F., Jamal, A., Bignell, E.M., Petrou, M.A., Coutts, R.H., 2012. Incidence of
549 dsRNA mycoviruses in a collection of Aspergillus fumigatus isolates. Mycopathologia
550 174, 323-326.

551 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden,
552 T.L., 2009. BLAST+: architecture and applications. BMC Bioinformatics 10, 421.

553 Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for
554 automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25,
555 1972-1973.

556 Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high
557 throughput. Nucleic Acids Res. 32, 1792-1797.

558 Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter,
559 S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., 2016. The Pfam protein families
560 database: towards a more sustainable future. Nucleic Acids Res. 44, D279-D285.

561 Ghabrial, S.A., Suzuki, N., 2009. Viruses of plant pathogenic fungi. Annu. Rev.
562 Phytopathol. 47, 353-384.
563 Gilbert, K.B., Holcomb, E.E., Allscheid, R.L., Carrington, J.C., 2019. Hiding in plain
564 sight: New virus genomes discovered via a systematic analysis of fungal public
565 transcriptomes. PLoS One 14, e0219207.
566 Grybchuk, D., Akopyants, N.S., Kostygov, A.Y., Konovalovas, A., Lye, L.F., Dobson,
567 D.E., Zangger, H., Fasel, N., Butenko, A., Frolov, A.O., Votypka, J., d'Avila-Levy,
568 C.M., Kulich, P., Moravcova, J., Plevka, P., Rogozin, I.B., Serva, S., Lukes, J.,
569 Beverley, S.M., Yurchenko, V., 2018. Viral discovery and diversity in trypanosomatid
570 protozoa with a focus on relatives of the human parasite Leishmania. Proc Natl Acad
571 Sci U S A 115, E506-E515.
572 Higashiura, T., Katoh, Y., Urayama, S., Hayashi, O., Aihara, M., Fukuhara, T., Fuji, S.I.,
573 Kobayashi, T., Hase, S., Arie, T., Teraoka, T., Komatsu, K., Moriyama, H., 2019.
574 Magnaporthe oryzae chrysovirus 1 strain D confers growth inhibition to the host
575 fungus and exhibits multiform viral structural proteins. Virology 535, 241-254.
576 Kartali, T, Nyilasi, I, Szabó, B, et al., 2019. Detection and Molecular Characterization
577 of Novel dsRNA Viruses Related to the *Totiviridae* Family in *Umbelopsis ramanniana*.
578 Front Cell Infect Microbiol 9: 249.
579 Khankhum, S., Escalante, C., de Souto, E.R., Valverde, R., 2017. Extraction and
580 electrophoretic analysis of large dsRNAs from desiccated plant tissues infected with
581 plant viruses and biotrophic fungi. Eur. J. Plant Pathol. 147, 431-441.
582 King, A.M.Q., Adams, M.J., Carstens, E.B., Lefkowitz, E.J., 2012. Virus Taxonomy:

583 Classification and Nomenclature of Viruses : Ninth Report of the International
584 Committee on Taxonomy of Viruses. Elsevier.
585 Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of
586 ribosomal RNAs in metatranscriptomic data. Bioinformatics 28, 3211-3217.
587 Lin, Y., Zhou, J., Zhou, X., Shuai, S., Zhou, R., An, H., Fang, S., Zhang, S., Deng, Q.,
588 2020. A novel narnavirus from the plant-pathogenic fungus Magnaporthe oryzae. Arch.
589 Virol. 165, 1235-1240.
590 Marzano, S.Y., Nelson, B.D., Ajayi-Oyetunde, O., Bradley, C.A., Hughes, T.J., Hartman,
591 G.L., Eastburn, D.M., Domier, L.L., 2016. Identification of Diverse Mycoviruses
592 through Metatranscriptomics Characterization of the Viromes of Five Major Fungal
593 Plant Pathogens. J. Virol. 90, 6846-6863.
594 Milne, I., Bayer, M., Cardle, L., Shaw, P., Stephen, G., Wright, F., Marshall, D., 2010.
595 Tablet—next generation sequence assembly visualization. Bioinformatics 26, 401-402.
596 Morris, T.J., Dodds, J.A., 1979. Isolation and Analysis of Double-Stranded-Rna from
597 Virus-Infected Plant and Fungal Tissue. Phytopathology 69, 854-858.
598 Okada, R., Kiyota, E., Moriyama, H., Fukuhara, T., Natsuaki, T., 2015. A simple and
599 rapid method to purify viral dsRNA from plant and fungal tissue. Journal of General
600 Plant Pathology 81, 103-107.
601 Rambaut, A., 2014. FigTree 1.4.2 software. Institute of Evolutionary Biology, Univ.
602 Edinburgh. URL http://tree.bio.ed.ac.uk/software/ figtree/.
603 Rastgou, M., Habibi, M.K., Izadpanah, K., Masenga, V., Milne, R.G., Wolf, Y.I., Koonin,
604 E.V., Turina, M., 2009. Molecular characterization of the plant virus genus

605 Ourmiavirus and evidence of inter-kingdom reassortment of viral genome segments as
606 its possible route of origin. J. Gen. Virol. 90, 2525-2535.
607 Refos, J.M., Vonk, A.G., Eadie, K., Lo-Ten-Foe, J.R., Verbrugh, H.A., van Diepeningen,
608 A.D., van de Sande, W.W., 2013. Double-stranded RNA mycovirus infection of
609 Aspergillus fumigatus is not dependent on the genetic make-up of the host. PLoS One
610 8, e77381.
611 Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and
612 post-analysis of large phylogenies. Bioinformatics 30, 1312-1313.
613 Sugiura, N., 1978. Further analysts of the data by akaike's information criterion and the
614 finite corrections. Communications in Statistics-Theory and Methods 7, 13-26.
615 Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular
616 evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725-2729.
617 Tanabe, A.S. (2011) Kakusan4 and Aminosan: two programs for comparing
618 nonpartitioned, proportional and separate models for combined molecular
619 phylogenetic analyses of multilocus sequence data. Mol Ecol Resour **11**: 914–921.
620 Turina, M, Ghignone, S, Astolfi, N, Silvestri, A, Bonfante, P, Lanfranco, L., 2018. The
621 virome of the arbuscular mycorrhizal fungus Gigaspora margarita reveals the first
622 report of DNA fragments corresponding to replicating non-retroviral RNA viruses in
623 fungi. Environ Microbiol 20: 2012–2025.
624 Urayama, S., Kato, S., Suzuki, Y., Aoki, N., Le, M.T., Arie, T., Teraoka, T., Fukuhara, T.,
625 Moriyama, H., 2010. Mycoviruses related to chrysovirus affect vegetative growth in
626 the rice blast fungus Magnaporthe oryzae. J. Gen. Virol. 91, 3085-3094.

627 Urayama, S., Katoh, Y., Fukuhara, T., Arie, T., Moriyama, H., Teraoka, T., 2014. Rapid
628 detection of Magnaporthe oryzae chrysovirus 1-A from fungal colonies on agar plates
629 and lesions of rice blast. Journal of General Plant Pathology 81, 97-102.
630 Urayama, S., Takaki, Y., Hagiwara, D., Nunoura, T., 2020. dsRNA-seq Reveals Novel
631 RNA Virus and Virus-Like Putative Complete Genome Sequences from Hymeniacidon
632 sp. Sponge. Microbes Environ. 35, ME19132.
633 Urayama, S., Takaki, Y., Nishi, S., Yoshida-Takashima, Y., Deguchi, S., Takai, K.,
634 Nunoura, T., 2018. Unveiling the RNA virosphere associated with marine
635 microorganisms. Mol Ecol Resour 18, 1444-1455.
636 Urayama, S., Takaki, Y., Nunoura, T., 2016. FLDS: A Comprehensive dsRNA
637 Sequencing Method for Intracellular RNA Virus Surveillance. Microbes Environ. 31,
638 33-40.
639 van Diepeningen, A.D., Debets, A.J., Hoekstra, R.F., 2006. Dynamics of dsRNA
640 mycoviruses in black Aspergillus populations. Fungal Genet. Biol. 43, 446-452.
641 Wolf, Y.I., Kazlauskas, D., Iranzo, J., Lucia-Sanz, A., Kuhn, J.H., Krupovic, M., Dolja,
642 V.V., Koonin, E.V., 2018. Origins and Evolution of the Global RNA Virome. mBio 9.
643 Zoll, J., Verweij, P.E., Melchers, W.J., 2018. Discovery and characterization of novel
644 Aspergillus fumigatus mycoviruses. PloS one 13.
645
646

**Table 1:** List of *Aspergillus* strains carrying viral sequences

| Host fungal species and strain | Virus name | Number of segments | Genome type [*4] | Detection method [*5] |
|---|---|---|---|---|
| ***A. fumigatus* IFM 62632** | Aspergillus fumigatus polymycovirus 1 (AfuPmV1) | 4 | dsRNA | AGE |
| | Aspergillus fumigatus negative-stranded RNA virus 1 (AfuNSRV1) [*1] | 1[*2] | ssRNA(-) | FLDS |
| ***A. fumigatus* IFM 63147** | Aspergillus fumigatus chrysovirus (AfuCV) | 4 | dsRNA | AGE |
| | Aspergillus fumigatus narnavirus 2 (AfuNV2) | 3 | ssRNA(+) | AGE |
| | Aspergillus fumigatus botourmiavirus 1 (AfuBOV1) [*1] | 1 | ssRNA(+) | FLDS |
| ***A. fumigatus* IFM 63431** | Aspergillus fumigatus narnavirus 2 (AfuNV2) | 3 | ssRNA(+) | AGE |
| ***A. fumigatus* IFM 64916** | Aspergillus fumigatus botourmiavirus 1 (AfuBOV1) [*1] | 1 | ssRNA(+) | AGE |
| ***A. fumigatus* IFM 62355** | Aspergillus fumigatus mitovirus 1 (AfuMV1) | 1 | ssRNA(+) | Pooled FLDS |
| ***A. fumigatus* IFM 62629** | Aspergillus fumigatus narnavirus 2 (AfuNV2) | 3 | ssRNA(+) | Pooled FLDS |
| ***A. fumigatus* IFM 63439** | Aspergillus fumigatus RNA virus 1 (AfuRV1) [*1] | 3 | ssRNA(+) | Pooled FLDS |
| ***A. fumigatus* IFM 64779** | Aspergillus fumigatus botourmiavirus 1 (AfuBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |
| ***A. lentulus* IFM 62627** | Aspergillus lentulus partitivirus 1 (AlePV1) [*1] | 2 | dsRNA | AGE |
| ***A. lentulus* IFM 63547** | Aspergillus lentulus narnavirus 1 (AleNV1) [*1] | 2 | ssRNA(+) | AGE |
| ***A. lentulus* IFM 64003** | Aspergillus lentulus non-segmented dsRNA virus 1 (AleNdsRV1) [*1] | 1 | dsRNA | AGE |
| ***A. lentulus* IFM 64004** | Aspergillus lentulus non-segmented dsRNA virus 1 (AleNdsRV1) [*1] | 1 | dsRNA | AGE |
| ***A. lentulus* IFM 65052** | Aspergillus lentulus totivirus 1 (AleTV1) [*1] | 1 | dsRNA | AGE |
| ***A. pseudoviridinutans* IFM 59502** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |
| ***A. pseudoviridinutans* IFM 59503** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1[*3] | ssRNA(+) | Pooled FLDS |
| ***A. pseudoviridinutans* IFM 61377** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |
| ***A. pseudoviridinutans* IFM 61378** | Aspergillus pseudoviridinutans botourmiavirus 1 (ApvBOV1) [*1] | 1 | ssRNA(+) | Pooled FLDS |

*1 : Tentative virus names were used to represent the novel sequences identified in the strains from this study.

*2 : The terminal sequence was not completely determined.

*3 : The presence of the viral sequence was confirmed by RT-PCR. However, the complete sequence was not determined.

*4 : The genome type of each novel virus was predicted based on that of similar viruses.

*5 : AGE, agarose gel electrophoresis; FLDS: the dsRNA from the single isolate was analyzed by FLDS; Pooled FLDS: the dsRNA samples from multiple (~20) isolates were pooled and analyzed by FLDS