

Discovery of new mycoviral genomes within publicly available fungal transcriptomic datasets

Kerrigan B. Gilbert¹, Emily E. Holcomb¹, Robyn L. Allscheid^{1,2}, James C. Carrington^{1*}

¹ Donald Danforth Plant Science Center, Saint Louis, Missouri, USA

² Current address: National Corn Growers Association, Chesterfield, Missouri, USA

* Address correspondence to James C. Carrington

E-mail: jcarrington@danforthcenter.org

Short title: Virus discovery from RNA-seq data

Abstract

The distribution and diversity of RNA viruses in fungi is incompletely understood due to the often cryptic nature of mycoviral infections and the focused study of primarily pathogenic and/or economically important fungi. As most viruses that are known to infect fungi possess either single-stranded or double-stranded RNA genomes, transcriptomic data provides the opportunity to query for viruses in diverse fungal samples without any *a priori* knowledge of virus infection. Here we describe a systematic survey of all transcriptomic datasets from fungi belonging to the subphylum Pezizomycotina. Using a simple but effective computational pipeline that uses reads discarded during normal RNA-seq analyses, followed by identification of a viral RNA-dependent RNA polymerase (RdRP) motif in *de novo* assembled contigs, 59 viruses from 44 different fungi were identified. Among the viruses identified, 89% were determined to be new species and 68% are, to our knowledge, the first virus described from the fungal species. Comprehensive analyses of both nucleotide and inferred protein sequences characterize the phylogenetic relationships between these viruses and the known set of mycoviral sequences and support the classification of up to four new families and two new genera. Thus the results provide a deeper understanding of the scope of mycoviral diversity while also increasing the distribution of fungal hosts. Further, this study demonstrates the suitability of analyzing RNA-seq data to facilitate rapid discovery of new viruses.

Introduction

Advances in low-cost, high-throughput sequencing technologies has revolutionized our ability to query the molecular world. Transcriptome projects are of particular use by providing targeted information about gene expression and improving the annotation of genome projects. Indeed, a number of “1K” transcriptome projects are underway, including the 1000 Plant Genome project, Fish-T1K project, 1K Insect Transcriptome Evolution (1KITE) project, and 1000 Fungal Genome and Transcriptome project. These large scale, collaborative projects have contributed to the vast amount of sequencing data available. As of February 2018, more than 16,000 petabases have been deposited in the Short Read Archive (SRA) at NCBI, with over 6,000 petabases available as open-access data.

Beyond revealing gene expression patterns within the target organism, these datasets are a source for additional insights as novel RNA sequences, specifically viral genomic sequences, have been identified within them. In an early example, analysis of RNA-seq data from a soybean cyst nematode revealed the presence of four different RNA viruses [1]. A more broad scale analysis of 66 non-angiosperm RNA-seq datasets, produced as part of the 1000 Plant Genome project, identified viral sequences in 28 plant species [2]. Additionally, the first viral sequences for any member of the family of spiders *Nephilidae* were identified through a secondary analysis of the RNA-seq data produced during the spider genome sequencing project [3].

While over 200 mycoviruses have been described, the full depth of the mycoviral landscape remains to be determined as certain biases have impacted the discovery of mycoviral sequences. Commonly infections of fungi appear cryptic, having little to no apparent impact on fungal health, thus occluding the presence of the virus. Additionally, many of the fungi studied are important human or plant pathogens, but this represents only a fraction of the known diversity of fungi. Advancements in mycovirus discoveries have been aided by screening the transcriptomes of various collections of fungi; study systems include those fungi associated with the seagrass *Posidonia oceanica* [4], five common plant-pathogenic fungal species [5], and soybean phyllosphere phytobiomes [6].

In the biosphere, fungi are not found in isolation; instead, they are members of complex environments that include members from some or all other kingdoms of life. Characterizing the responses and contributions that fungi make to these environments also includes a full understanding of the dynamics at play within the fungi. We hypothesized that mycoviral

infections are more pervasive than currently known, and as such, undertook a systematic approach to identify new mycoviral species by querying the publicly available fungal transcriptomic datasets at the Short Read Archive hosted at NCBI. Our bioinformatic approach provides a robust method for the identification of new viral species from raw sequencing data as we identified 59 viral genomes, 53 of which are described here for the first time.

Results and Discussion

Viral RNA-dependent RNA polymerases (RdRP) are essential enzymes in the genomes of RNA viruses that have no DNA stage. The domain organization of this protein is described as a right hand, where three subdomains are the palm, fingers, and thumb [7]. While some of the motifs found within these subdomains are conserved among all polymerase enzymes, specific structures and motifs are found only in RNA-dependent RNA polymerases. In addition to unique signatures within the sequences themselves, virus-like RdRPs are not encoded in the DNA genomes of eukaryotic organisms, thus viral RdRPs can be specifically targeted as a marker for sequence-based analyses.

We previously determined that contigs assembled from fungal RNA-seq data can be queried for viral RdRP-specific domains [8]. As such, we undertook a systematic analysis of all RNA-seq datasets from Pezizomycotina fungi (division: Ascomycota; subkingdom: Dikarya) available at the Short Read Archive (SRA) at the National Center for Biotechnology Information (NCBI) on July 21, 2017. An initial list of 1,127 unique BioProjects was identified, which was manually curated to identify 569 BioProjects suitable for this analysis (S1 Table). Samples were excluded for a variety of reasons including: individually listed BioProjects were determined to belong to larger groups of experiments using the same fungal strain and were therefore combined and a single representative sample was selected; samples containing a mix of fungi and RNA from other kingdoms were excluded as the host of any putative viral sequences could not be determined; samples where the read sequences were too short for assembly or unavailable for download.

Virus-like RdRP Discovery Pipeline

A computational pipeline was created to automate the steps following selection of a raw reads SRA file available and identification of putative viral RdRP sequences. The majority of the steps used publicly available programs in conjunction with a small number of custom Perl scripts. The

Table 1: Summary of mycoviral genomes identified from fungal RNA-seq datasets

Fungal Host	BioProject	SRA ID(s)	Data Reference ^a	Mapped Read	Read Coverage	Segment Number	Length (nt)	Nucleotide Accession	Protein ID	Length (aa)
Acidomyces richmondensis BFW	PRJNA250470	SRR1797521	JGI (SWS)	80,068	1,167	1	10291	MK279511	ORF1-RdRP	2182
									ORF3	859
									ORF4	334
Aspergillus ellipticus strain CBS 707.79	PRJNA250911	SRR1799565	JGI (SEB)	806,965	19,358	1	6253	MK279500	RdRP	1556
									ORF2	482
Aspergillus heteromorphus isolate CBS 117.55	PRJNA250969	SRR1799577	JGI (SEB)	662,010	27,769	1	3576	MK279437	RdRP	1124
				654,402	35,799	2	2742	MK279438	Coat Protein	833
				1,306,352	80,739	3	2427	MK279439	Hyp. Protein	727
Aspergillus homomorphus strain CBS 101889	PRJNA250984	SRR1799578	JGI (SEB)	288	8	1	5147	MK279489	Coat Protein	780
									RdRP	826
Aspergillus homomorphus strain CBS 101889	PRJNA250984	SRR1799578	JGI (SEB)	3,944	163	1	3621	MK279487	RdRP	963
Aspergillus neoniger isolate CBS 115656	PRJNA250996	SRR1801411	JGI (SEB)	18,188	1,007	1	2708	MK279481	RdRP	769
Beauveria bassiana strain ARSEF 8028	PRJNA260878	SRR3269778	Valero-Jiménez et al. 2016	8,635	245	1	3170	MK279499	ORF1	315
									RdRP	590
Beauveria bassiana strain GXsk1011	PRJNA306902	SRR3043102	Wang et al. 2017	2,761	71	1	3478	MK279433	RdRP	1115
				356	10	2	3143	MK279434	Coat Protein	976
				860	25	3	3069	MK279435	Hyp. Protein	964
				2,190	71	4	2770	MK279436	Hyp. Protein	851
Clohesyomyces aquaticus strain CBS 115471	PRJNA372822	SRR5494002	Mondo et al. 2017	22,928	1,848	1	1873	MK279448	RdRP	578
				29,817	2,434	2	1850	MK279449	Coat Protein	536
Colletotrichum caudatum strain CBS 131602	PRJNA262373	SRR5145569	JGI (JC)	941	28	1	5125	MK279490	Coat Protein	787
									RdRP	825
Colletotrichum eremochloae isolate CBS 129661	PRJNA262412	SRR5166050	JGI (JC)	938	78	1	1815	MK279452	RdRP	573
				3,764	333	2	1696	MK279453	Coat Protein	494
Colletotrichum eremochloae isolate CBS129661	PRJNA262412	SRR5166050	JGI (JC)	1,312	37	1	5364	MK279491	Coat Protein	757
									RdRP	862
Colletotrichum falcatum strain CoC 671	PRJNA272832	SRR1765657	JGI (JC)	10,585	464	1	2283	MK279482	RdRP	703
Colletotrichum navitas strain VF38c	PRJNA262225	SRR5166338	JGI (JC)	1,666	49	1	5070	MK279492	Coat Protein	757
									RdRP	831
Colletotrichum zoysiae strain MAFF235873	PRJNA262216	SRR5166062	JGI (JC)	18,146	526	1	5174	MK279493	Coat Protein	783
									RdRP	831
Cryphonectria parasitica strain cpku80	PRJNA369604	SRR5235483	Nerva et al. 2017	315	18	1	1800	MK279462	RdRP	573
				90	5	2	1670	MK279463	Coat Protein	494
Delitschia confertaspora strain ATCC 74209	PRJNA250740	SRR3440303	JGI (JS)	2,798	221	1	1899	MK279446	RdRP	571
				16,760	1,528	2	1645	MK279447	Coat Protein	501
Drechslerella stenobrocha strain 248	PRJNA236481	SRR1145648	Liu et al., 2014	4,529	234	1	1935	MK279440	RdRP	586
				12,254	712	2	1721	MK279441	Coat Protein	521
Fusarium graminearum strain HN10	PRJNA263651	SRR4445678	Wang et al. 2016	323,426	3,102	1	13035	MK279472	ORFA	176
									ORFB	3705
Fusarium poae isolate 2516	PRJNA319914	SRR3953125	Vanheule et al. 2016	8,168	363	1	2271	MK279442	RdRP	701
				128,392	5,897	2	2199	MK279443	Coat Protein	637
Gaeumannomyces tritici strain GGT-007	PRJNA268052	SRR1664730	Yang et al. 2015	2,145,797	33,888	1	6332	MK279501	RdRP	1525
				ORF2	504					
Gaeumannomyces tritici strain GGT-007	PRJNA268052	SRR1664730	Yang et al. 2015	940	50	1	1881	MK279464	RdRP	578
				130,207	7,487	2	1739	MK279465	Coat Protein	494
Gaeumannomyces tritici strain GGT-007	PRJNA268052	SRR1664730	Yang et al. 2015	4,639	225	1	2058	MK279466	RdRP	617
				860	48	2	1798	MK279467	Coat Protein	487
Grosmanina clavigera strain KW1407	PRJNA184372	SRR636711	Wang et al. 2013	97,131	4,761	1	2040	MK279468	RdRP	599
				478,553	28,101	2	1703	MK279469	Coat Protein	479

Fungal Host	BioProject	SRA ID(s)	Data Reference ^a	Mapped Read	Read Coverage	Segment Number	Length (nt)	Nucleotide Accession	Protein ID	Length (aa)
Gyromitra esculenta strain CBS 101906	PRJNA372840	SRR5491178	JGI (JS)	4,977	52	1	14584	MK279476	Polyprotein	4846
Hortaea werneckii strain EXF-2000	PRJNA356640	SRR5086622	Sinha et al. 2017	12,385	341	1	4580	MK279498	Coat Protein RdRP	681 805
Loramycetes juncicola strain ATCC 46458	PRJNA372853	SRR5487643	JGI (JS)	8,698	540	1	2416	MK279483	RdRP	688
Magnaporthe grisea strain M82	PRJNA269089	SRR1695911	Luo et al. 2015	314 199	28 20	1 2	1725 1530	MK279458 MK279459	RdRP Coat Protein	539 437
Morchella importuna strain SCYDJ1-A1	PRJNA372858	SRR5487664	JGI (JS)	3,146	29	1	16495	MK279477	Polyprotein	5447
Morchella importuna strain SCYDJ1-A1	PRJNA372858	SRR5487664	JGI (JS)	4,264	42	1	15394	MK279478	Polyprotein	4956
Morchella importuna strain SCYDJ1-A1	PRJNA372858	SRR5487664	JGI (JS)	4,814	51	1	14296	MK279479	Polyprotein	4734
Morchella importuna strain SCYDJ1-A1	PRJNA372858	SRR5487664	JGI (JS)	6,933	134	1	7835	MK279502	RdRP ORF2 ORF3	1480 381 622
Morchella importuna strain SCYDJ1-A1	PRJNA372858	SRR5487664	JGI (JS)	14,742	220	1	10099	MK279480	Polyprotein	3295
Neurospora discreta strain FGSC8579	PRJNA257829	SRR1539773	Lehr et al. 2014	66,037	755	1	6648	MK279503	RdRP ORF2	1526 522
Ophiocordyceps sinensis strain 1229	PRJNA292632	SRR2533613	Li et al. 2016	1,219	51	1	2386	MK279484	RdRP	698
Ophiocordyceps sinensis strain 1229	PRJNA292632	SRR2533613	Li et al. 2016	3,114	126	1	2439	MK279485	RdRP	716
Penicillium brasilianum strain MG11	PRJEB7514	ERR677271	Hom et al. 2015	2,358 4,222	132 278	1 2	1791 1518	MK279470 MK279471	RdRP Coat Protein	569 453
Penicillium digitatum KH8	PRJNA254400	SRR1557148	Wang et al. 2015	1,395	27	1	5184	MK279494	Coat Protein RdRP	807 838
Penicillium digitatum KH8	PRJNA254400	SRR1557148	Wang et al. 2015	3,209	62	1	5168	MK279495	Coat Protein RdRP	776 825
Penicillium digitatum strain HS-F6	PRJNA352307	SRR4851221	CCNU (QC)	1,070 1,903	60 125	1 2	1781 1526	MK279456 MK279457	RdRP Coat Protein	539 434
Penicillium digitatum strain HS-F6	PRJNA352307	SRR4851221	CCNU (QC)	813	23	1	3507	MK279488	RdRP	987
Penicillium raistrickii strain ATCC 10490	PRJNA250734	SRR1801283	JGI (DS)	182 166 206 400	8 8 10 20	1 2 3 4	3501 3120 2955 2980	MK279429 MK279430 MK279431 MK279432	RdRP Coat Protein Hyp. Protein Hyp. Protein	1116 983 901 847
Periconia macrospinoso strain DSE2036	PRJNA262386	SRR5153210	Knapp et al. 2018	11,576	487	1	3566	MK279507	ORF1 (HP) RdRP	336 459
Phyllosticta citriasiana strain CBS 120486	PRJNA250394	SRR3314103	JGI (JS)	136 110	12 11	1 2	1700 1504	MK279450 MK279451	RdRP Coat Protein	535 421
Pseudogymnoascus destructans strain 20631-21	PRJNA66121	SRR254216	Broad Institute	391 277	17 13	1 2	1754 1562	MK279444 MK279445	RdRP Coat Protein	539 434
Rutstroemia firma strain CBS 115.86	PRJNA372878	SRR5689333	JGI (JS)	7,749,083	175,028	1	6641	MK279504	RdRP ORF2	1558 545
Sclerotinia homoeocarpa isolate LT11	PRJNA167556	SRR515157	Hulvey et al. 2012	12,790	180	1	7171	MK279505	RdRP ORF2	1718 589
Sclerotinia homoeocarpa isolate LT11	PRJNA167556	SRR515157	Hulvey et al. 2012	7,913	65	1	12370	MK279473	Polyprotein	3604
Setosphaeria turcica strain 28A	PRJNA250530	SRR1587420	CU (GT)	1,222	53	1	3478	MK279508	ORF1 (HP) RdRP	344 496
Setosphaeria turcica strain 28A	PRJNA250530	SRR1587420	CU (GT)	19,494	322	1	9069	MK279474	Polyprotein	2751
Setosphaeria turcica strain 28A	PRJNA250530	SRR1587420	CU (GT)	1,082	63	1	2595	MK279486	RdRP	652
Thelebolus microsporus ATCC 90970	PRJNA372886	SRR5487623	PNNL (JM)	208 2,476	18 243	1 2	1696 1531	MK279460 MK279461	RdRP Coat Protein	533 444

Fungal Host	BioProject	SRA ID(s)	Data Reference ^a	Mapped Read	Read Coverage	Segment Number	Length (nt)	Nucleotide Accession	Protein ID	Length (aa)
Thelebolus microsporus strain ATCC 90970	PRJNA372886	SRR5487623	PNNL (JM)	604	18	1	5146	MK279496	Coat Protein	763
									RdRP	829
Tolypocladium ophioglossoides strain CBS 100239	PRJNA292830	SRR2179765	Quandt et al. 2016	47,721	408	1	5264	MK279497	Coat Protein	787
									RdRP	831
Trichoderma asperellum strain CBS 131938	PRJNA261111	SRR1575447	Bech et al. 2015	103,958	658	1	14211	MK279475	Polyprotein	4176
Trichoderma citrinoviride strain FP-102208	PRJNA304029	SRR2961293	Lin et al. 2017	372	16	1	2284	MK279454	RdRP	727
				217,293	9,554	2	2297	MK279455	Coat Protein	663
Trichoderma harzianum strain TR274	PRJNA216008	SRR976276, SRR976277, SRR976278	Steindorff et al. 2014	1,289	33	1	3911	MK279509	ORF1 (HP)	464
									RdRP	493
Verticillium longisporum isolate 43	PRJNA308558	SRR3102592	EMAUG (KH)	402	15	1	2798	MK279510	ORF1 (HP)	256
									RdRP	494
Zymoseptoria tritici strain IPO323	PRJNA179083	SRR612175	UE (DS)	258,267	3,548				RdRP	1491
	PRJEB8798	ERR789231	Rudd et al. 2015	480,473	8,049	1	5969	MK279506		
	PRJNA237967	SRR1167717	Kellner et al. 2014	1,823,937	15,278					
	PRJNA253135	SRR1427070	Omrane et al. 2015	851,723	7,277				ORF2	499

^a Abbreviations corresponding to the contact person for unpublished datasets:

JGI (SWS): Joint Genome Institute, Steven W. Singer

JGI (SEB): Joint Genome Institute, Scott E. Baker

JGI (JC): Joint Genome Institute, Jo Ann Crouch

JGI (JS): Joint Genome Institute, Joseph Spatafora

CCNU (QC): Central China Normal University, Qianwen Cao

JGI (DG): Joint Genome Institute, Dave Greenshields

CU (GT): Cornell University, Gillian Turgeon

PNNL (JM): Pacific Northwest National Laboratory, Jon Magnuson

EMAUG (KH): Ernst-Moritz-Arndt University of Greifswald, Katharina Hoff

UE (DS): University of Exeter, David Studholme

overall structure of the pipeline was optimized to efficiently work within a HTCondor-managed system including the use of the DAGMan (Directed Acyclic Graph Manager) meta-scheduler to coordinate the required dependencies between pipeline steps.

Due to the highly specific nature of viral RdRP protein sequences, the key feature of this pipeline was the use of a customized Pfam database which included the viral RdRP families: RdRP_1, RdRP_2, RdRP_3, RdRP_4, RdRP_5, and Mitovir_RNA_pol. To err on the side of over-inclusiveness, HMMER *hmmScan* was run with an e-value cut-off of 10.0, allowing for the identification of any sequence with limited similarity to the viral protein domains. Putative RdRP sequence(s) were identified in a total of 1067 contigs from 284 of the 569 samples. The actual number of viruses was expected to be less than this total number; reasons include a false positive match to the viral RdRP signatures, contigs that are the forward and reverse strand sequence of a single virus, or contigs that are partial matches along a single virus sequence. To confirm similarity to known viral sequences, contigs were evaluated via a BLASTX search versus the 'nr' database at NCBI followed by manual evaluation and curation of contigs with hits to known mycoviral sequences.

Summary of identified viral sequences

In total, 59 complete, RNA mycoviral genomes were identified in 47 of the 569 SRA samples analyzed. A global analysis of the sequencing data available and the resultant virus sequences along with the set of known mycoviruses available at Virus-HostDB (version March 23, 2018, <http://www.genome.jp/virushostdb>) demonstrates a number of patterns and trends (Fig 1). There is an uneven distribution of data being generated within the Pezizomycotina subphylum whereby certain Classes are more abundantly represented than others (illustrated by the width of the band going from "Fungal Class" to "Fungal Genus"). For instance, the Sordariomycetes and Eurotiomycetes host 77% of the sequencing projects while the remaining eight Classes host the other 23% of datasets. Unsurprisingly, fungi known to be key pathogens of plants and humans are present within the Sordariomycetes and Eurotiomycetes, including the well-studied genera *Fusarium*, *Aspergillus*, and *Penicillium*. Indeed, seven genera alone account for 51% of all sequencing projects analyzed in this study, with the remaining 49% of datasets spread among 176 other genera.

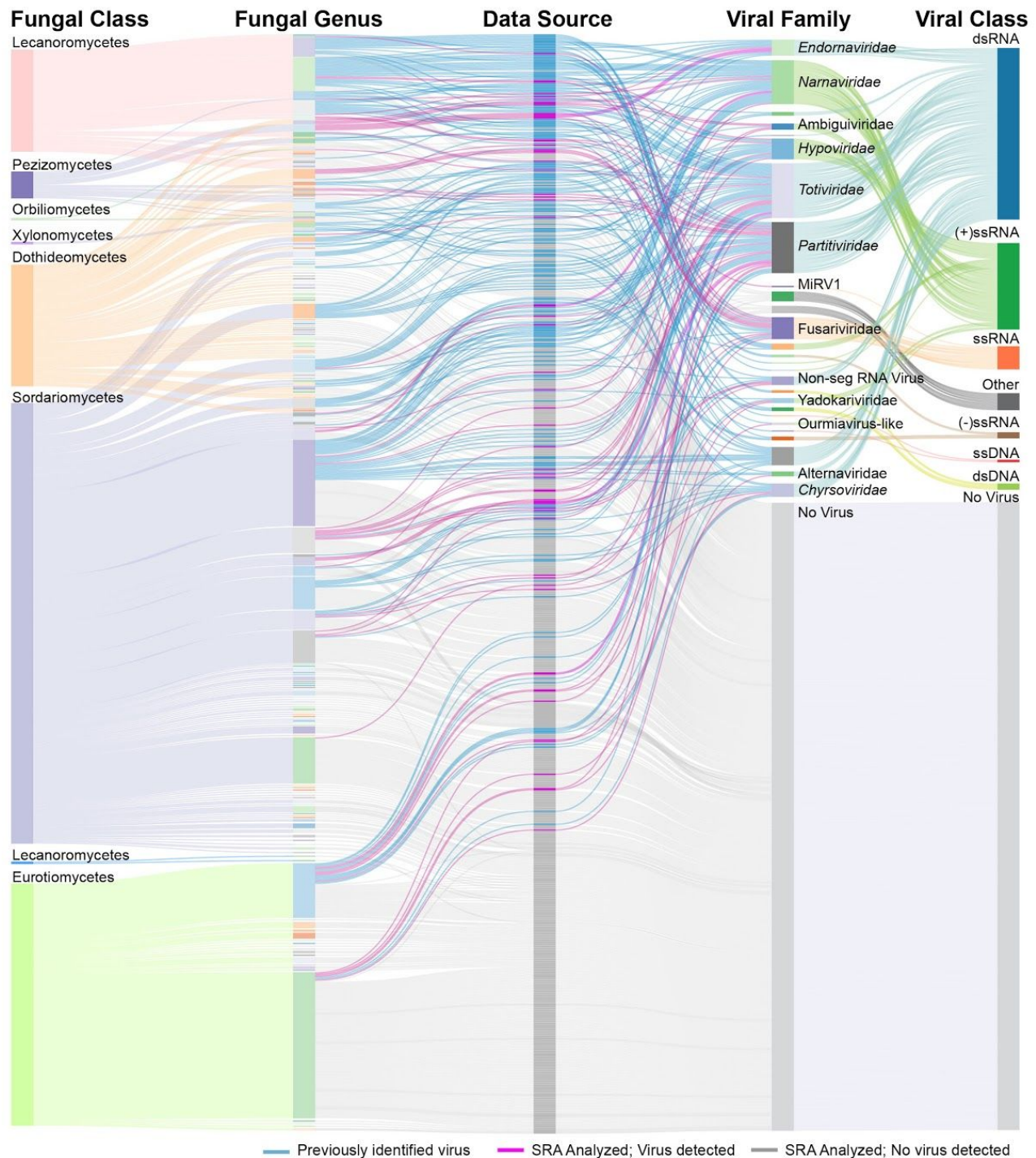


Fig 1. Sankey diagram summarizing both the SRA samples queried in this study and recognized mycoviral sequences. The height of the bars in each column is indicative of the number of samples. Classification of fungal host is summarized in the two, leftmost columns. Column "Fungal Class": fungal species are sorted into one of eight taxonomic classes, all genera from the same class are the same color; Column "Fungal Genus": further separates the fungi into their respective genera; colors used in this column are to highlight the individual genera but similar colors have no relation to one another. The third column, "Data Source" is color-coded by the origin of the sample where blue indicates a known mycovirus, while pink and grey indicate an SRA sample was analyzed and either a virus sequence was or was not detected. The two rightmost columns summarize the information related to virus taxonomy. Column 4: "Viral Family" organizes the samples with either a known or discovered virus by viral family. The names of families and groups that contain a virus identified in this study are indicated. Column 5: "Viral Class" summarizes the viral genomes by nucleic acid and/or strandedness. Samples without a virus detected are in the category "No Virus".

Virus discovery as a function of number of sequencing projects favors the less data-rich Classes of fungi. While Sordariomycetes and Eurotiomycetes host 61% of the viruses identified in this study, the discovery rate is 17%. Conversely, the less-studied classes Leotiomycetes, and Pezizomycetes had virus discovery rates of 21%, and 26% respectively. Clearly, however, a certain threshold of data is required for virus discovery, as the Classes Xylonomycetes, Orbiliomycetes, and Lecanoromycetes had a combined total of seven sequencing projects but only one virus was identified.

Comparing the distribution of the known set of Pezizomycotina-hosted mycoviruses with those identified in this study demonstrates similar distribution patterns. Members of the Sordariomycetes and Eurotiomycetes Classes are host to 58% of known mycoviruses, with Leotiomycetes being also abundant. Gains in under-represented Classes are observed, as only two mycoviruses were previously known to be hosted by a Pezizomycetes fungus, and five viruses were identified in this study. Additionally, the first virus from the Class Orbiliomycetes was found in this study.

Classification of the virus sequences identified in this study reveals that 19% of viruses are from unclassified viral families. This is similar to the known set of mycoviruses, where 18% of viruses are also unclassified. Further, dsRNA viruses are the most abundant category within the two viral datasets, with 63% and 54% belonging to those identified in this study and the known set respectively. The ratio of dsRNA to ssRNA viruses in this study is higher than the known dataset, indicating an under-identification of ssRNA viruses (18% versus 27%, this study and known dataset respectively). This result may indicate that the RNA preparation methods for RNA-seq experiments are biased against ssRNA viruses.

An in-depth analysis and characterization of each of the 59 viral sequences is described below (Table 1). Viral genomes have been divided into two major categories: dsRNA viruses and (+)ssRNA viruses then further separated by viral family and, where applicable, genus. In total, viruses were identified from eight of the currently recognized virus families and as such are described within the context of the expected characteristics. The remaining viruses form groups with other recently identified mycoviruses leading to the characterization of new mycoviral families and genera within an existing family.

Double stranded RNA (dsRNA) Viruses

A total of 34 viruses were categorized as dsRNA viruses due the identity of the top BLASTP match. To further characterize the relationship of these sequences, a phylogenetic tree based on the RdRP amino acid sequences from the 34 identified viruses and the top BLASTP matches was constructed (Fig 2). Globally, five groups are visible, where three are known viral families (*Totiviridae*, *Partitiviridae*, and *Chrysoviridae*) and two represent potentially new families. Properties characteristic of each group of viruses are described in detail below.

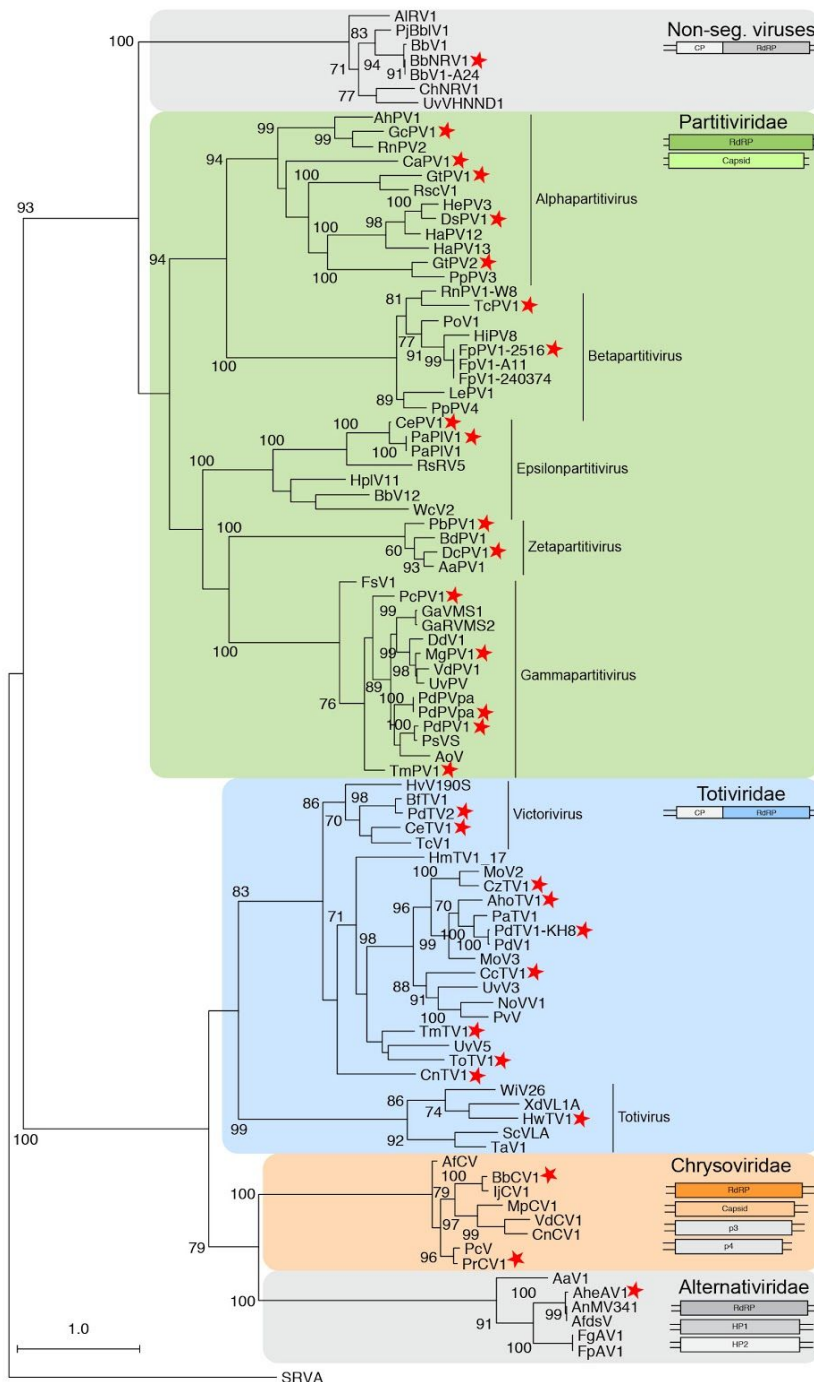


Fig 2. Phylogram of the RdRP-containing ORF for viruses classified as dsRNA viruses. The sequence from *Simian Rotavirus A* (SRVA) serves as the outgroup. Three officially recognized families are present: *Partitiviridae*, *Chrysoviridae*, and *Totiviridae* and are highlighted with a green, orange, or blue box respectively. Gray boxes surround families currently unrecognized. The structure of the genomic segment(s) for each highlighted group is depicted below the family name. Genera within the *Partitiviridae* and *Totiviridae* family are indicated. Viruses identified in this study are noted with a red star. Numbers at the nodes indicate bootstrap support over 70% (1000 replicates). RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein; HP: Hypothetical Protein.

Non-segmented RNA Virus (BbNRV1)

Recently, dsRNA viruses have been identified from various fungi that encode two ORFs within the single-segment genome, where the RdRP sequences have more similarity to the RdRPs of partitiviruses (multi-segment genomes) than totiviruses (mono-segment genomes). The virus *Beauveria bassiana non-segmented dsRNA virus 1* (BbNRV1), assembled using the RNA-seq data from a transcriptome experiment of the highly virulent isolate *B. bassiana* ARSEF 8028 [9], groups with these previously identified viruses (Fig 2). A second virus from this group was also identified in this analysis: *Colletotrichum higginsianum non-segmented RNA virus 1* (ChNRV1; this virus served as an internal control for the virus discovery pipeline as it was previously identified via a similar approach from two different *Colletotrichum higginsianum* RNA-seq datasets [8].

BbNRV1 and ChNRV1, together with four previously identified viruses, form a strongly supported group that is separate from other characterized fungal viruses (Fig 2). It has been noted by multiple researchers that these fungal viruses share similarities with a family of plant viruses, the *Amalgaviridae*, specifically the mono-segmented genome structure and partitivirus-like RdRP sequence. Therefore, a phylogenetic analysis of the RdRP protein sequence was performed which demonstrates that, while the mono-segmented genome structure is shared, this new clade of fungal viruses is also distinct from the plant amalgaviruses (Fig 3A). Indeed, the amalgavirus group of plant viruses branches off before the BbNRV1-related and *Partitivirus* clades. A percent identity matrix generated with RdRP and CP coding sequences further demonstrates a separation between this new group and existing virus families (Fig 3B). For the RdRP pairwise comparisons, 50% identity easily differentiates the three clades from each other. Additionally, the CP sequence comparisons demonstrate shared similarities within the three clades and less than 20% similarity is shared between the clades.

The genome structure of viruses within this group consists of two predicted ORFs surrounded by variable length UTRs (Fig 3C). The 5'UTR lengths fall into two distinct ranges: less than 40 nt for *Penicillium janczewskii Beauveria bassiana-like virus 1* (PjBbIV1) [4], ChNRV1 [8], and UvHNND1 [10], or between 260 to 320 nt for the three viruses isolated from *B. bassiana* [11,12] and *Alternaria longipes dsRNA virus 1* (AIV1-HN28) [13]. The protein encoded by ORF1 is predicted to be from 314 to 394 aa and in a +1 frame relative to the protein encoded by ORF2. The intergenic sequence lengths are generally less than 100 nt, with the exception of AIV1-HN28. ORF2 nucleotide sequences range from 1578 to 1773 nt, and encoded proteins, which encode an RdRP domain, range in size from 525 to 590 aa in length. A slippery site

sequence identified between ORF1 and ORF2 of ChNRV1 suggested that an ORF1-ORF2 fusion protein may be made via a -1 ribosomal frameshift; indeed, trypsin digest followed by mass spectrometry of a protein near the predicted weight of the fusion protein returned signatures for both the ORF1 and ORF2 protein sequences [8]. Numerous attempts to isolate virus particles definitely produced by a virus within this group remain unsuccessful [4,8,11].

Due to the shared genome structure and partitivirus-like RdRP protein sequence, early viruses from this group were tentatively grouped with the plant *Amalgaviridae* family. However, with the identification of additional genome sequences, a more in depth phylogenetic analysis demonstrates the mycoviruses described here are separate. A new genus, “Unirnavirus”, has recently been suggested to reflect the single, unified genome

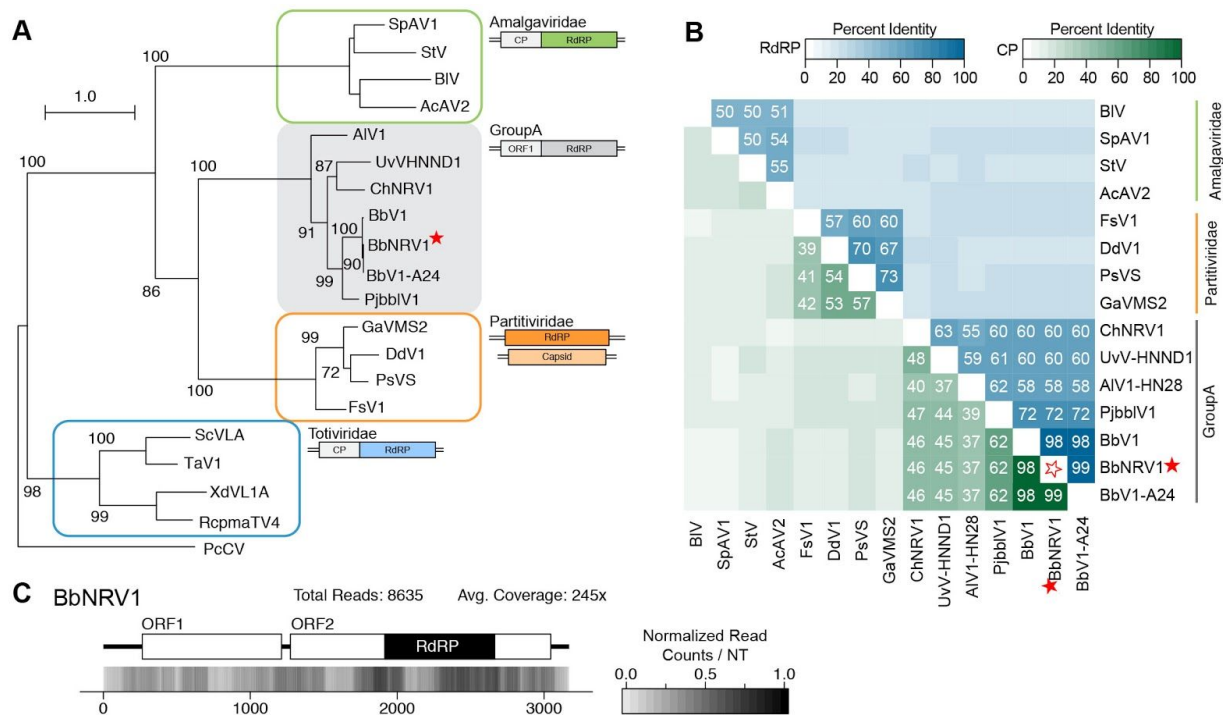


Fig 3. Analyses of identified viral genomes in relation to known *Amalgaviridae* and *Partitiviridae*. (A) Phylogram of the ORF containing the RdRP motif for members of the new mycovirus family along with select partitiviruses, totiviruses and amalgaviruses; sequence from *Penicillium chrysogenum virus* (*PcCV*) serves as the outgroup. Full names and accession numbers for protein sequences used in the alignment are in Table S2. Scale bar on the phylogenetic tree represents 1.0 amino acid substitutions per site and numbers at the nodes indicate bootstrap support over 70% (1000 replicates). Diagrams illustrating the genome structure and predicted ORFs are present for each group of viruses. (B) Percent identity matrix, generated by Clustal-Omega 2.1, of new mycoviral family, partitiviruses, and amalgaviruses. The top half of the matrix, in the blue scale, is the percent identity of the RdRP protein coding sequence and the bottom half of the matrix, in the green scale, is the percent identity of the putative CP. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the virus identified. Where the identified viral sequences have significant similarity to known viral sequences is indicated, specifically RdRP sequences with similarity $\geq 50\%$ and CP sequences with similarity $\geq 30\%$. (C) Density plot of read counts per nucleotide across the viral genome for *BbNRV1*. Graphical depiction of the genome structure, ORFs and predicted motifs are on along the top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. The red star indicates the virus identified in this study. RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein; ORF: Open Reading Frame.

segment that contains both ORFs [12]. As percent identity is a common criteria for species demarcation, the pairwise comparisons here provide possible thresholds for determining new species within this group of viruses; excluding the three viruses isolated from different strains of *B. bassiana*, a threshold of 75% identity in the RdRP protein sequence and 70% in the ORF1 coding sequence would be appropriate criteria to distinguish between the viruses isolated from unique fungal hosts.

Alternaviridae

Initially, a single segment encoding a single gene with an RdRP domain was identified from an *Aspergillus heteromorphus* sequencing sample (PRJNA250969; unpublished). BLAST alignment to 'nr' identified five related RdRP mycoviral sequences from viruses with multi-segmented genomes. *Fusarium graminearum alternavirus 1* (FgAV1) includes at least one additional genomic segment while *Alternaria alternata virus 1* (AaV1) [14] and *Aspergillus foetidus dsRNA mycovirus* (AfdsV) [15] each contain a total of four genomic segments. Using the additional sequences from these related viruses, two more contigs were identified from the *Aspergillus heteromorphus* Trinity assembly. Thus, the AheAV1 genome contains at least three genomic segments, the largest of which encodes the RdRP enzyme and the two smaller segments that encode proteins of unknown function (Table 1). Recently it has been proposed that a new family be established, "*Alternaviridae*" [15] we support the formation of this family, and the requisite genus (*Alternavirus*) and propose that AheAV1 is also a member of this new mycoviral family.

Previously it was observed that members of this clade group together, but separate from other families [14,15]. A phylogenetic analysis using the RdRP coding sequence demonstrates that the group containing AheAV1 forms a well supported clade that appears related to the *Chrysoviridae* and *Totiviridae* families (Fig 2). Members of the Totiviridae family have single-segment genomes containing two ORFs, while members of the Chrysoviridae have multi-segment genomes. While the RdRP segment sizes are similar between alternaviruses and chrysoviruses (~3,500 bp), the remaining two to three segments are hundreds of base pairs shorter than those of chrysoviruses. Further, the proteins of unknown function have no predicted domains or shared sequence homology outside of this group of viruses. The three viral RdRP sequences isolated from different species of *Aspergillus*, including AheAV1, are most closely related; similarly, the RdRP sequences from the two viruses from *Fusarium* species are most related (Fig 2). Indeed, a pairwise percent identity comparison of both the RdRP and HP1

protein sequences demonstrates the very high identity among the viruses from the two *Fusarium* and also from the two *Aspergillus* viruses (Fig 4A). This analysis also illustrates a clear delineation between alternaviruses and chrysovirus.

A common feature to viral genomes with multiple segments is the presence of a conserved sequence within the 5' UTRs of each fragment. Sequence conservation was described within the UTRs of the other putative alternaviruses, including AfV [15] and AaV1 [14]. Alignment of the 5'UTR sequences of the three genomic segments of AheAV1 demonstrates sequence

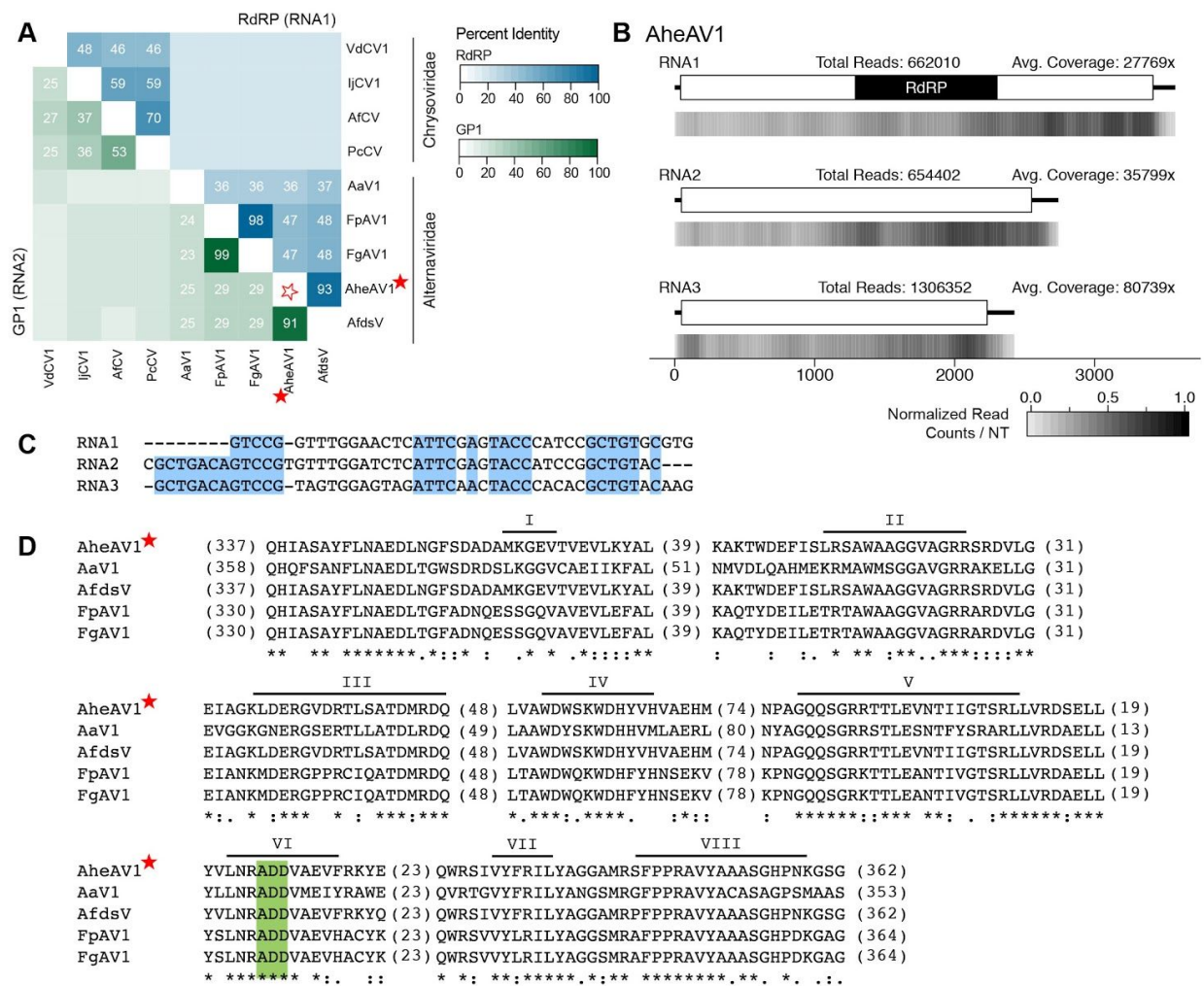


Fig 4. Analyses of *Aspergillus heteromorphus alternavirus 1* (AheAV1), a member of the recently proposed family *Alternaviridae*. (A) Percent identity matrix, generated by Clustal-Omega 2.1, of alternaviruses and members of the *Chrysoviriidae* family. The top half of the matrix, in the blue scale, is the percent identity of the RdRP protein coding sequence (from the RNA1 segment) and the bottom half of the matrix, in the green scale, is the percent identity of the second largest genomic fragment, encoding GP1. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the virus identified. Values $\geq 20\%$ identity are noted. (B) Density plot of read counts per nucleotide across the viral genomic segments of AheAV1. Graphical depiction of the genome structure, predicted ORFs and motifs for each segment are on top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. (C) Alignment of 5'UTR sequences of the three genomic segments of AheAV1. Conserved nucleotides are highlighted in blue. (D) Alignment of RdRP domains of Alternaviruses found in (A). The eight conserved RdRP motifs of dsRNA polymerases are noted along the top of the alignment and the ADD triad, found only in Alternavirus sequences, is highlighted in green. The number of amino acids not shown in alignment are noted in parentheses. Viruses identified in this study are indicated with a red star. RdRP: RNA-dependent RNA Polymerase; GP1: Gene Product.

conservation (Fig 4C). Alignment of the source RNA-seq reads to the viral genome segments reveals a high amount of total reads, with over four million reads mapped across the three segments (Fig 4B). This represents nearly 25% of the reads used for the analysis (reads that did not align to the fungal genome) and over 5% of the total reads from the sample. In an effort to identify a possible fourth genomic fragment for AheAV1, both the 5'UTR conserved sequence and the high read count were used as criteria to further query the Trinity assembled contigs however no additional contigs were identified.

An alignment of the RdRP protein sequences from the five alternavirus genomes reveals the presence of the expected eight conserved domains found within viral RdRP proteins (Fig 4D). Certain differences are shared among these viruses that are unique when compared to known viral sequences. Specifically, the triad within domain VI has an alanine (ADD) instead of the nearly universally conserved glycine (GDD).

Taken together, a number of criteria identify a new *Alternavirus*: multiple genome segments, where the largest encodes an RdRP protein, and the presence of an alanine within motif VI of the RdRP protein sequence. Classification of a new species within the genus *Alternavirus* remains to be established. To date, the distinguishing factors between viruses within this genus are the fungal host species and the number of genomic segments. Using these criteria, AheAV1 is a new species within the genus *Alternavirus*.

Totiviridae

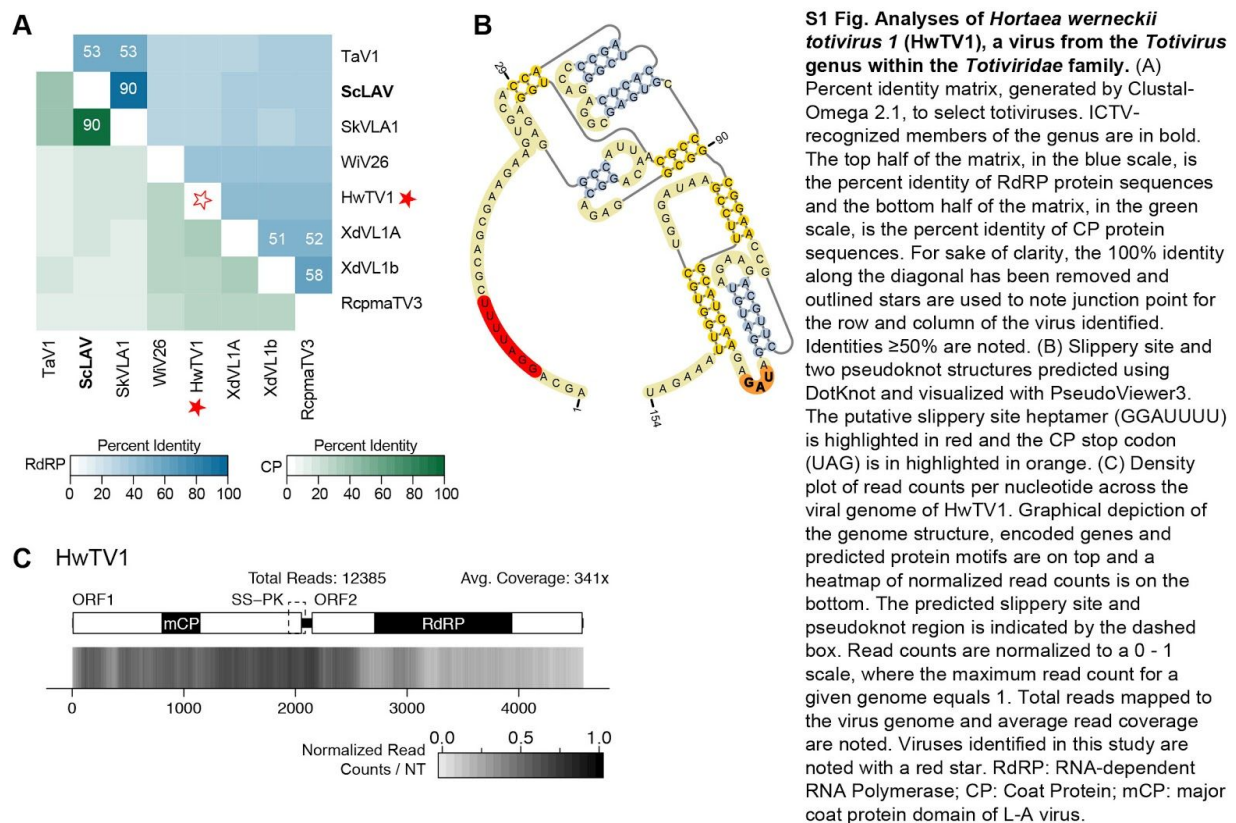
Ten of the viral sequences identified in this study are similar to other known members of the *Totiviridae* family, nine within the genus *Victorivirus* and one within the genus *Totivirus*.

Totiviruses

Hortaea werneckii totivirus 1 (HwTV1), the first viral sequence described for this fungal species to our knowledge (PRJNA356640; [16]), is a member of the *Totivirus* genus (Fig 2). Species demarcation within *Totivirus* requires biological characterization of the virus within a host cell; however, isolation of a virus from a distinct host species and less than 50% sequence identity of the RdRP at the protein level is a proxy for identifying a separate species [17]. HwTV1 was identified from a new fungal host, *H. werneckii*, and shares less than 44% sequence similarity to the most closely related known totiviruses (S1A Fig) thus satisfying the proxy criteria for a new virus species. The RdRP protein sequence of HwTV1 shares the greatest similarity to two totiviruses identified within the fungus *Xanthophyllum dendrorhous*, XdV-L1-A and

XdV-L1-B [18], along with the invertebrate virus *Wuhan insect virus 26* (WiV26), isolated from RNA-sequencing data of a mixed sample of flea and ants isolated in Hubei, China [19].

A common feature to members of the *Totivirus* genus is the presence of two partially overlapping open reading frames that produce a fusion protein via a -1 frameshift event with the 5' ORF encoding the major capsid protein and the 3' ORF encoding an RdRP. Characteristics of a frameshift site are (1) a heptanucleotide sequence 'slippery site' with the sequence X XXY YYZ and (2) a downstream RNA pseudoknot [20]. As expected, the two open reading frames in HwTV1 are frame 3 (CP) and frame 2 (RdRP). A canonical slippery site, G GGU UUU, is present upstream of the ORF1 stop codon at position 1,940 - 1,946 nt, and two predicted pseudoknots are present immediately 3' of the slippery site (S1B Fig). A conserved domain search within the encoded proteins reveals the presence of the major coat protein (L-A virus) domain in ORF1, in addition to the RdRP domain in ORF2 (S1C Fig) A total of 12,385 reads align to the HwTV1 genome, resulting in an average coverage of 341x, with slightly higher coverage observed across the 5' half of the genome (S1C Fig).

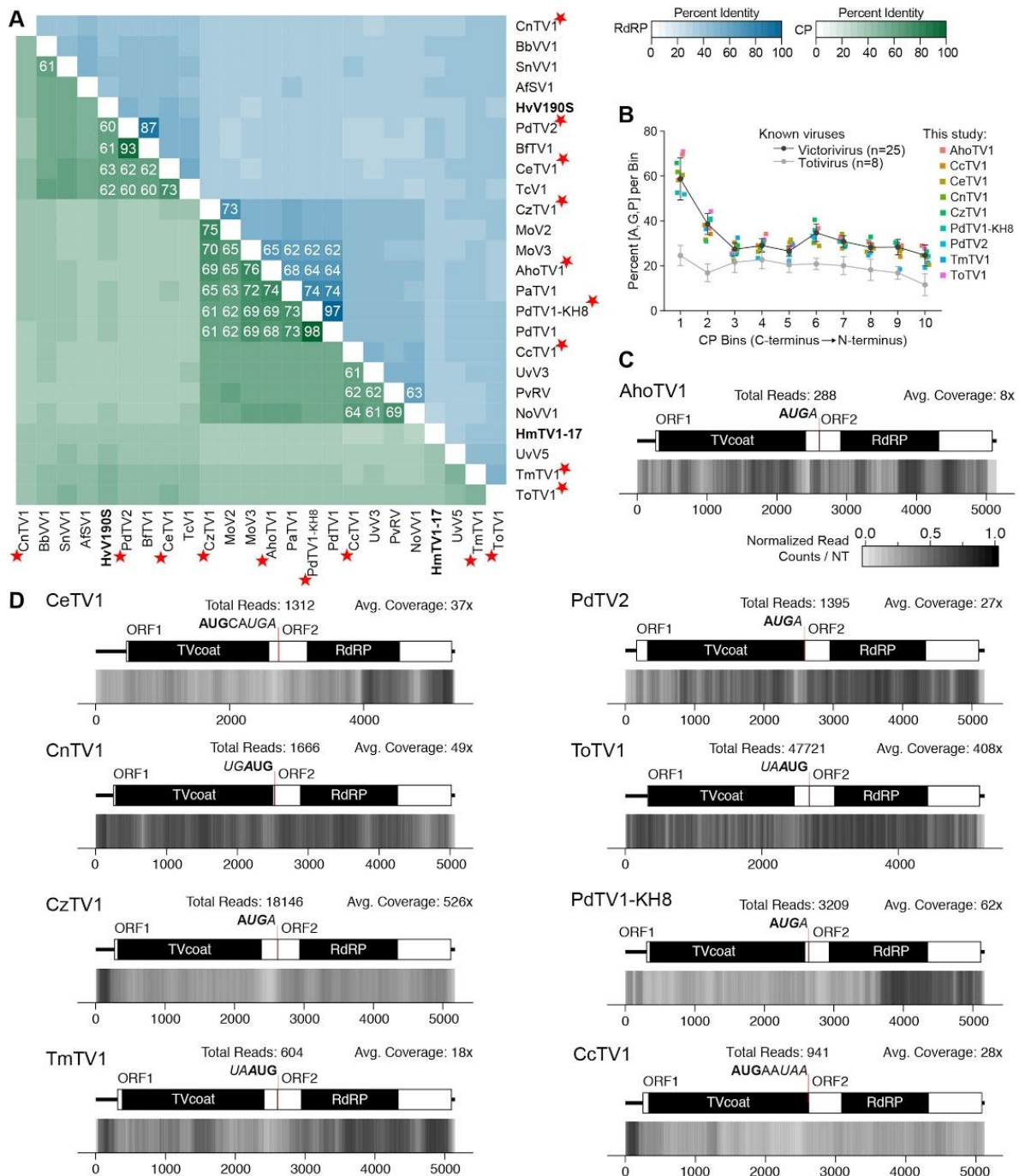


Victoriviruses

The remaining nine viruses in the *Totiviridae* family were found within two branches of the *Victorivirus* genus. The prototype virus of this genus, *Helminthosporium victoriae virus 190S* (HvV190S) [21], is present within one clade along with *Colletotrichum eremochloae totivirus 1* (CeTV1) (PRJNA262412; unpublished), *Colletotrichum navitas totivirus 1* (CnTV1) (PRJNA262225; unpublished), and *Penicillium digitatum totivirus 2* (PdTV2) (PRJNA254400; [22]). A second clade contains *Colletotrichum caudatum totivirus 1* (CcTV1) (PRJNA262373; unpublished), *Colletotrichum zoysiae totivirus 1* (CzTV1) (PRJNA262216; unpublished), *Penicillium digitatum totivirus 1-KH8* (PdTV1-KH8) (PRJNA254400; [22]), *Thelebolus microsporus totivirus 1* (TmTV1) (PRJNA372886; unpublished), *Tolypocladium ophioglossoides totivirus 1* (ToTV1) (PRJNA292830; [23]), and *Aspergillus homomorphus totivirus 1* (AhoTV1) (PRJNA250984; unpublished) along with other victoriviruses (Fig 2).

Alignment of the RNA-seq reads to the viral genomes reveals a wide range of average coverage values: from 8x to 526x (S2C and S2D Fig). A peak in coverage observed in CcTV1 and CzTV1 immediately precedes an octanucleotide sequence (AGGGUCC) which has been observed in the 5'UTR of some other victoriviruses including *Alternaria arborescens victorivirus 1* (AaV1) [24] and *Epichoë festucae virus 1* (EfV1) [25]. The peak in coverage of CnTV1 immediately precedes an octanucleotide sequence slightly different than previously reported (CGGCCUCC). A similar peak in coverage was not observed for AhoTV1, which also contains the AAGGGUCC octamer sequence within the 5'UTR, although this viral sequence had the lowest average coverage. The genome structure diagrams drawn above the coverage heat maps highlight features of the viruses that are further described below.

Unlike the frame-shifting observed in totiviruses, the RdRP of Victoriviruses is expressed as a separate protein from the coat protein [26] via a termination-reinitiation strategy [27]. The AUG start codon either overlaps the stop codon of the upstream of ORF1 or slightly precedes it, and is nearly always found in the -1 frame relative to ORF1 [21]. The tetranucleotide sequence, **AUGA**, is observed for PdTV2, PdTV3, CzTV1, and AspTV1. For CnTV1, TmTV1, and ToTV1, the last nucleotide of the ORF1 stop codon overlaps the ORF2 start codon, creating the pentamer sequences **UGAUG**, **UAAUG**, and **UAAUG** respectively. In each of these seven instances, there is the expected -1 frameshift between ORF1 and ORF2. For



S2 Fig. Analyses of identified viral genomes belonging to the *Victorivirus* genus within the *Totiviridae* family. (A) Percent identity matrix, generated by Clustal-Omega 2.1, comparing identified viral sequences to select members of the genus *Victorivirus*. The top half of the matrix, in the blue scale, is the percent identity of RdRP protein sequences and the bottom half of the matrix, in the green scale, is the percent identity of CP protein sequences. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the viruses identified. Where the identified viral sequences have significant similarity to known viral sequences is indicated, specifically RdRP or CP sequences with similarity $\geq 60\%$. Virus names in bold are ICTV-recognized species within the genus. (B) Percent of alanine, glycine, and proline (A,G,P) amino acids in the CP coding sequence. The average values were calculated for twenty-five sequences from known victoriviruses and ten sequences from known totiviruses, shown as dark and light grey lines respectively. The nine putative victoriviruses identified in this study were individually analyzed and plotted. (C and D) Density plot of read counts per nucleotide across the viral genome for the nine viruses identified. Graphical depiction of the genome structure, predicted ORFs, and putative motifs are illustrated on top and a heatmap of normalized read counts is on the bottom. Reads are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted for each virus. The overlap between the CP stop codon (italics) and RdRP start codon (bolded) is also indicated. Red stars indicate viruses identified in this study. RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein; ORF: Open Reading Frame.

CeTV1 and CcTV1, a two nucleotide spacer is observed between the start codon of ORF2 and the stop codon of ORF1 (**AUGCAUGA** and **AUGAAUAA** respectively) that is similar to the sequences observed in SsRV1 and SsRV2 (**AUGAAUAA** and **AUGAGUAA** respectively) [28]. Thus ORF2 of CeTV1 and CcTV1, as well as SsRV1 and SsRV2, is found in the +1 frame relative to ORF1.

A characteristic of the coat protein sequences of victoriviruses is an abundance of alanine, glycine, and proline (A,G,P) residues, particularly towards the C-terminal end. An analysis of between the CP protein sequences from victorivirus species and totivirus species illustrates a generally elevated percentage of these three amino acids across the entire CP coding sequence of victoriviruses in comparison to totiviruses, with a notable increase observed in the two bins closest to the C-terminus of the protein (S2B Fig). The same pattern is also observed for all nine proposed victorivirus sequences identified in this study.

These analyses demonstrate that these nine sequences share conserved characteristics found in members of the genus *Victorivirus*. Criteria for species demarcation within this genus are fungal host species and percent identity of the RdRP and CP amino acid sequences [17], with a cutoff of 60% identity for both. The CP and RdRP sequences of CnTV1, TmTV1, and ToTV1 satisfy both the sequence-based criteria of < 60% identity and unique fungal host criteria to be characterized as new species (S2A Fig).

For a growing number of recently published sequences with strong phylogenetic ties to victoriviruses, the cutoff of 60% identity for the RdRP and/or CP is not satisfied, only the criteria of a differing fungal host (S2A Fig). For five of the six other Victoriviruses identified in this study, a 60% pairwise identity would not distinguish between viral sequences, only the host-based criteria. The sixth virus, PdTV1-KH8, shares 97% (RdRP) and 98% (CP) identity with the already published sequence for *Penicillium digitatum virus 1* (PdV1) [29] thus neither the sequence-based nor host-based criteria are met. Notably, PdTV2 and PdTV1-KH8 were both identified from the same data sample, however they share only 33.5% identity. The coincident infection of *P. digitatum* KH8 by two victoriviruses is not unprecedented: *Sphaeropsis sapinea RNA virus 1* and *Sphaeropsis sapinea RNA virus 2* were isolated from the same host and share less than 60% similarity [28].

Clearly a precise sequence-based cut-off to distinguish between mycoviral species has not yet been identified. Until further biological studies to query the host range of these similar viruses can be performed, we would recommend more stringent values for pairwise percent identities for species demarcation within the *Victorivirus* genus: 80% for CP sequences and 90%

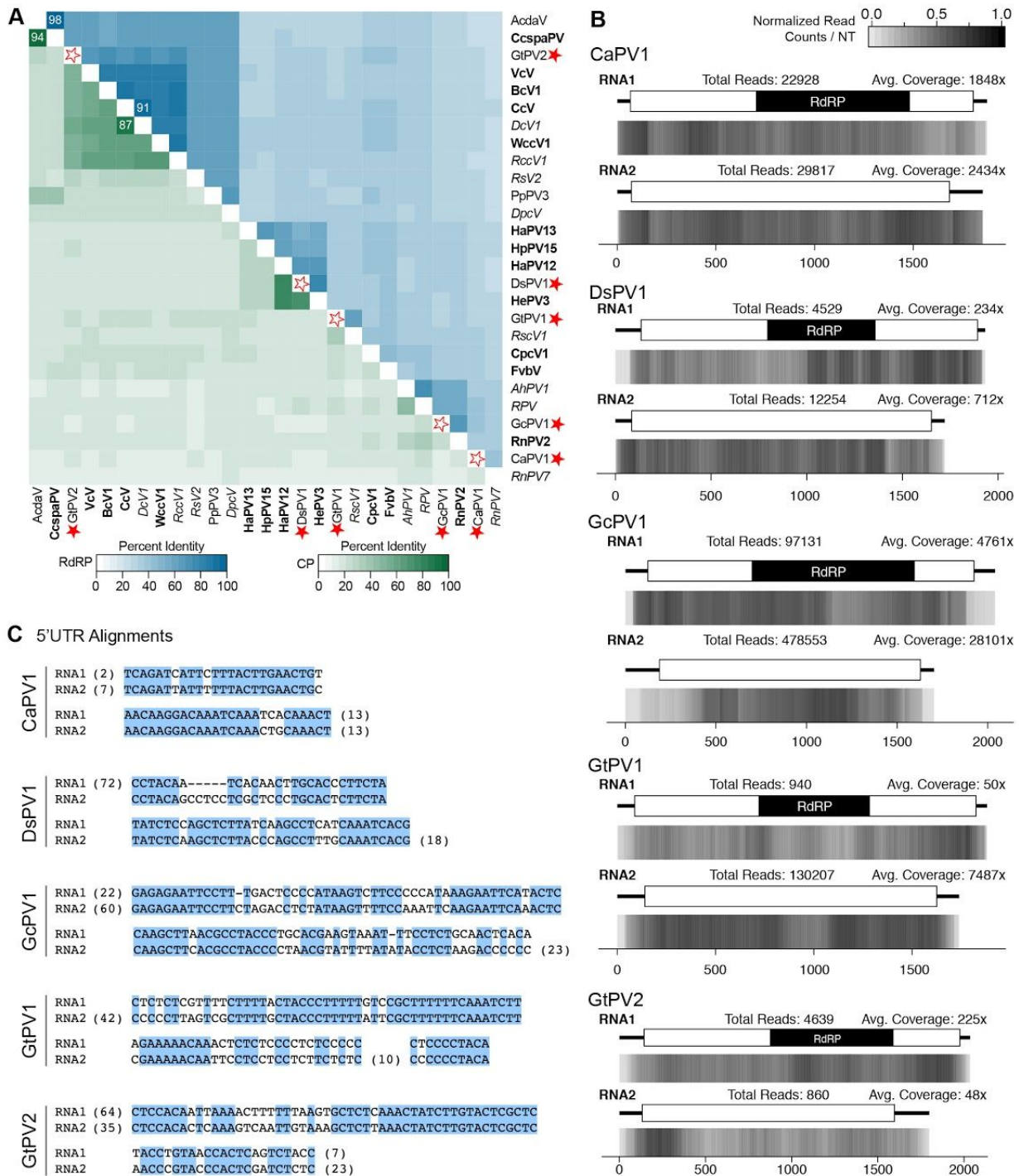
for RdRP sequences in addition to the identity of the fungal host. Following this recommendation, the eight viruses, CnTV1, TdTV2, CeTV1, CcTV1, CzTV1, AspTV1, TmTV1, and ToTV1, would be identified as new viral species within the *Victorivirus* genus. To our knowledge, this is the first mycovirus described for *C. navitas*, *C. eremochloae*, *C. zoysiae*, *C. caudatum*, *Thelebolus microsporus*, and *Tolypocladium ophioglossoides*.

Partitiviridae

Five genera have been assigned to the family *Partitiviridae*, along with fifteen viruses unassigned to a genus [30]. Genome structure of this family requires two distinct genome segments, each containing a single ORF: dsRNA1 and dsRNA2. The RdRP is found on dsRNA1 while the CP is found on dsRNA2, and highly conserved sequences can be found in the 5'UTR of both viral genome segments [30]. To date, viruses isolated from fungi have been classified into one of three of the *Partitiviridae* genera: *Alphapartitivirus*, *Betapartitivirus*, and *Gammapartitivirus*. In this study, sixteen viruses were identified that clearly group among the known partitiviruses (Fig 2).

Alphapartitivirus

Five viruses identified in this study phylogenetically group within the *Alphapartitivirus* genus: *Clohesyomyces aquaticus partitivirus 1* (CaPV1) (PRJNA372822; [31]), *Grosmannia clavigera partitivirus 1* (GcPV1) (PRJNA184372; [32]), *Drechlerella stenobrocha partitivirus 1* (DsPV1) (; PRJNA236481[33]), *Gaeumannomyces tritici partitivirus 1* and 2 (GtPV1, GtPV2) (PRJNA268052; [34]). To our knowledge, the first three are the first viruses identified in these fungal species while the two partitiviruses in *G. tritici* are the first partitiviruses from this fungus. As with other members of this genus, the genomic segment encoding the RdRP is approximately 1.9 to 2.0 kb while the genomic segment encoding the CP is between 1.7 and 1.9 kb (Table 1). Alignment of the RNA-seq reads to the viral genome sequences reveals average coverage per nucleotide ranging from 48x to 28,101x (S3B Fig). In general, the two dsRNA segments do not appear to have similar coverage per nucleotide across the segments, and in all but one instance, the dsRNA2 segment has the higher average coverage. For the two viruses isolated from *G. tritici*, 23x more reads align to GtPV1 than GtPV2 (S3B Fig).



S3 Fig. Analyses of viral genomes for members of the genus *Alphapartitivirus*. (A) Percent identity matrix generated by Clustal-Omega 2.1 of viruses identified in this study and select alphapartitiviruses. The top half of the matrix, in the blue scale, is the percent identity of RdRp protein sequences and the bottom half of the matrix, in the green scale, is the percent identity of CP protein sequences. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note the junction for the row and column of the viruses identified. Identities above the species demarcation are noted: $\geq 90\%$ for RdRp and $>80\%$ for CP. Names in bold are ICTV-member species, while names in italics are related, unclassified viruses. (B) Density plot of read counts per nucleotide across the viral genome for the five alphapartitiviruses identified in this study. Graphical depiction of the genome structures of RNA1 and RNA2, the predicted ORFs and RdRP motif are on illustrated above the heatmap of normalized read counts. Reads are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted for each virus. (C) Alignment of 5'UTRs of RNA1 and RNA2 genome segments. Residues not shown in alignment are in parentheses. Blue shading indicates regions of 100% identity. Red stars note the viruses identified in this study. RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein.

Multiple sequence alignments of the 5'UTR regions of the two RNA genomic sequences for each identified viral genome revealed highly conserved stretches of nucleotides specific to each virus (S3C Fig). As dsRNA satellite sequences have been previously described for the alphapartitivirus *Cherry chlorotic rusty spot associated partitivirus* [35] the conserved sequences present in dsRNA1 and dsRNA2 were used to query the remaining Trinity contigs, however no putative satellite RNAs were identified.

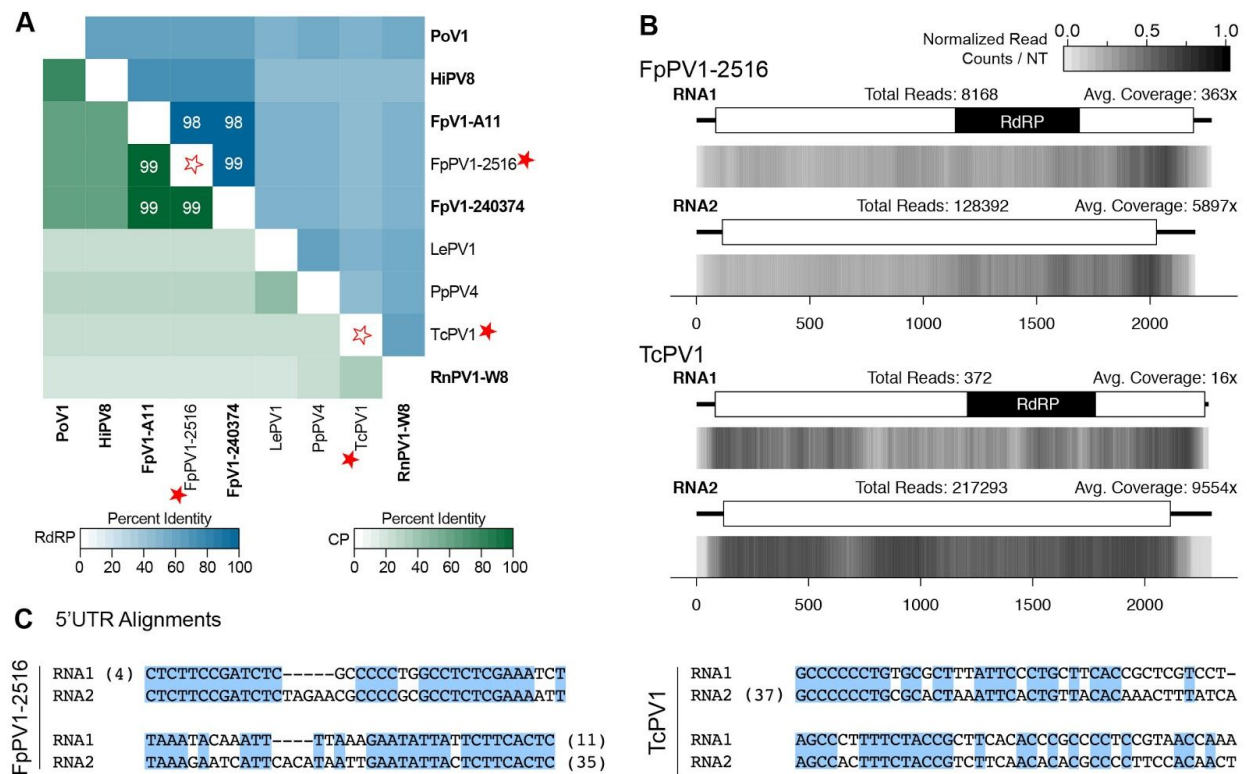
The criteria for species demarcation within the genus *Alphapartitivirus* are $\leq 90\%$ identity in the RdRP protein sequence and $\leq 80\%$ identity in the CP protein sequence. All five viral genomes identified in this study satisfy these criteria and as such are new species of the genus *Alphapartitivirus* (S3A Fig).

Betapartitivirus

Two virus genomes were identified as members of the genus *Betapartitivirus*: *Trichoderma citrinoviride partitivirus 1* (TcPV1) (PRJNA304029; [36]) and *Fusarium poae partitivirus 1-2516* (FpPV1-2516) (PRJNA319914; [37]). To our knowledge, this is the first partitivirus identified from *T. citrinoviride* while FpPV1-2516 is the third instance of a species of betapartitivirus described to date from the fungus *F. poae*.

The RdRP coding sequence from FpPV1-2516 shares 99.9% and 97.9% identity with FpV1 (*F. poae* strain A11) and FpV1 (*F. poae* strain 240374) respectively (S4A Fig). Similarly, the CP sequence of FpPV1-2516 shares 99.8% identity with both of the FpV1-CP sequences. Despite the high sequence similarity observed, these three viruses appear to originate from different locales. FpV1-240374 was published in 2016 following deep RNA-sequencing of the fungal strain *F. poae* MAFF 240374, a fungus isolated from a *Triticum aestivum* spikelet in Japan in 1991 [38]. FpV1(strain A11) was identified in 1998 from the strain *F. poae* A-11 isolated from wheat grain in Hungary [39]. Finally, FpPV1-2516, identified in this study, was found in the RNA-seq data from *F. poae* strain 2516, a fungus isolated in Belgium from wheat [37]. While FpPV1-2516 is not a unique viral species, it does have a unique genome sequence and unique history from the other two known viruses and as such, has been deposited at Genbank (Table 1). A pairwise percent identity analysis of TcPV1 with other betapartitiviruses reveals that this is a unique sequence since identity values for both the RdRP and CP sequences are below the established thresholds for species demarcation of 90% identity and 80% respectively (S4A Fig) [30].

A sequence alignment of the 5'UTR sequences of the two genomic segments of TcPV1 reveals stretches of highly conserved nucleotides, characteristic of partitiviral genomes (S4C Fig). Additionally, the two genomic segments are both 2.3 kb (Table 1), which is within the described range for betapartitiviruses of 2.1 to 2.4 kb [30]. As observed with the majority of the alphapartitiviruses described above, the average coverage per nucleotide for the genomic segments of FpPV1-2516 and TcPV1 was higher for the dsRNA2 segment that encodes the CP than the RdRP segment (S4B Fig).



S4 Fig. Analyses of identified viral genomes belonging to the genus *Betapartitivirus*. (A) Percent identity matrix, generated by Clustal-Omega 2.1, between identified viruses and select betapartitiviruses. The top half of the matrix, in the blue scale, is the percent identity of RdRP protein sequences and the bottom half of the matrix, in the green scale, is the percent identity of CP protein sequences. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the viruses identified. Where the percent similarity is above the cutoffs for species demarcation is noted, specifically >90% for RdRP sequences and >80% for CP sequences. Bolded names are ICTV-recognized member species. Full names and accession numbers for protein sequences are in Table S2. (B) Density plot of read counts per nucleotide across the viral genome for the two *Betapartitiviruses*. Graphical depiction of the genome structure of the two RNAs, the predicted ORFs, and the RdRP motif are along the top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted for each virus. (C) Alignment of 5'UTR sequences from the RNA1 and RNA2 genome segments. Blue boxes highlight conserved nucleotides and numbers in parentheses are residues not shown. Viruses identified in this study are noted with a red star. RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein.

Gammapartitivirus

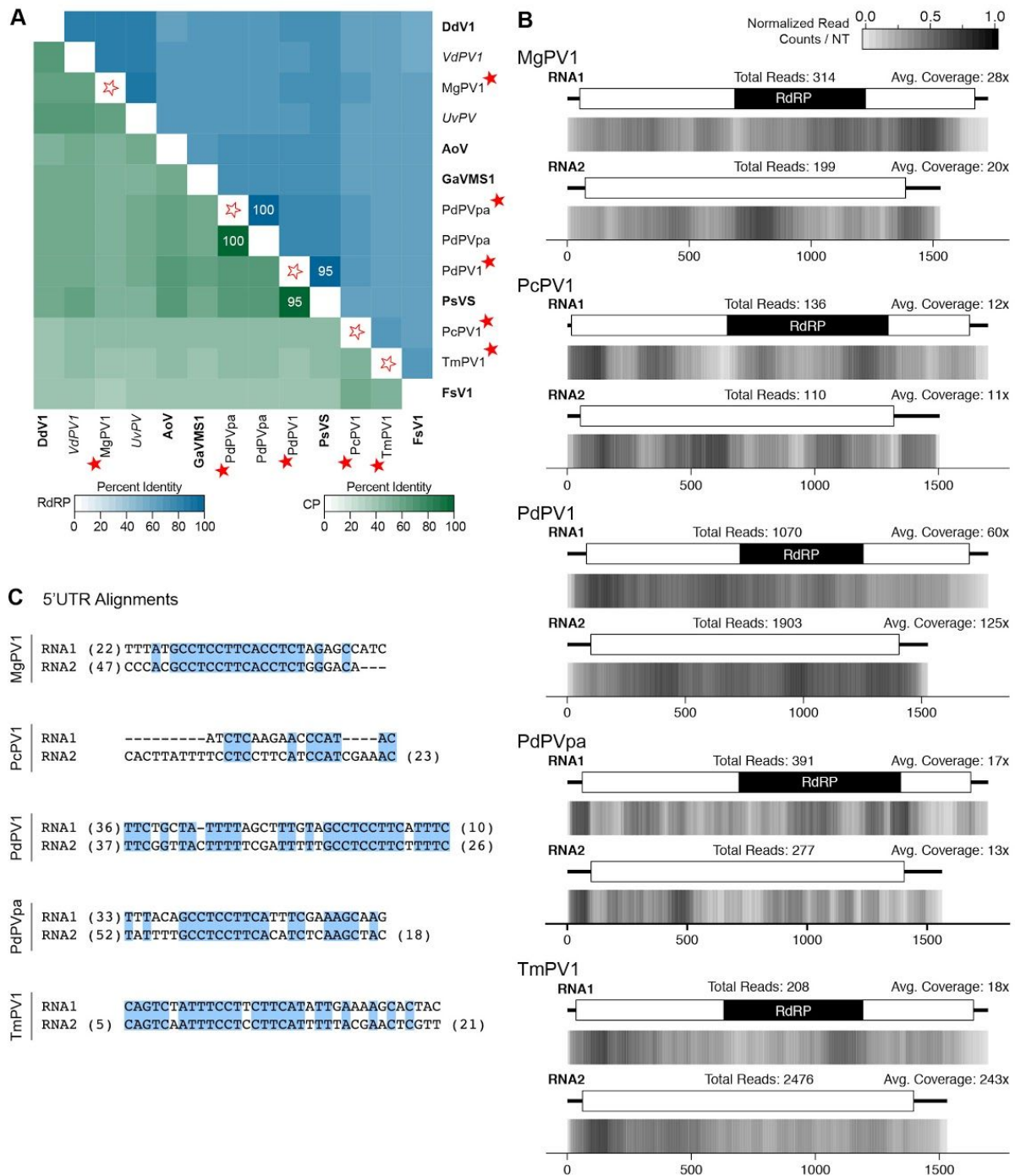
Five viruses identified in this analysis grouped with members of the genus *Gammapartitivirus* (Fig 2): *Penicillium digitatum partitivirus 1-HSF6* (PdPV1-HSF6) (PRJNA352307; unpublished)

and *Pseudogymnoascus destructans partitivirus 1-20631* (PdPV1-20631) (PRJNA66121; unpublished) along with *Magnaporthe grisea partitivirus 1* (MgPV1) (PRJNA269089; [40]), *Thelebolus microsporus partitivirus 1* (TmPV1) (PRJNA372886; unpublished), and *Phyllosticta citriasiana partitivirus 1* (PcPV1) (PRJNA250394; unpublished). The latter three viruses are, to our knowledge, the first mycoviruses described for these fungal species and, as described below, satisfy the criteria to be identified as new viral species.

The dsRNA1 genomic segment encoding the RdRP for all five viruses is approximately 1.7 kb while the dsRNA2 genomic segment encoding the CP is between 1.5 and 1.6 bp (Table 1) both within the expected range of other known members of the genus *Gammapartitivirus* of 1.6 to 1.8 kb for dsRNA1 and 1.4 to 1.6 kb for dsRNA2 [30]. A multiple sequence alignment analysis of the 5'UTR sequences from the two genomic segments of each virus revealed stretches of conserved nucleotides within four of the five viruses; PcPV1 was the exception, however the 5'UTR sequence of dsRNA1 is only 17 nt (S5C Fig). Attempts to extend this sequence by manually querying the RNA-seq data were unsuccessful.

Total reads aligned and average coverage per nucleotide were calculated following alignment of the RNA-seq reads to the viral genome sequences. In general, the average coverage per nucleotide for dsRNA1 is similar to that of dsRNA2 (S5B Fig), a departure from the alignment results for the alphapartitiviruses and betapartitiviruses described above.

A pairwise percent identity analysis compared the RdRP coding sequences and the CP coding sequences for the five viruses to other sequences within the genus *Gammapartitivirus* (S5A Fig). The current criteria for species demarcation is 90% for the RdRP sequence and 80% for the CP sequence [30]. Genic and protein sequences from PdPV1-20631 share 100% identity with the virus PdPVpa from *Pseudogymnoascus destructans* 80251, although differences were observed in the length of the 5' and 3' UTR sequences. The RdRP and CP protein sequences of PdPV1-HSF6, isolated from *Penicillium digitatum*, shares 95% identity with the proteins from *Penicillium stoloniferum virus S* (PsVS). Due to the high sequence similarity and the relative close relationship between the two fungal hosts from the *Penicillium* genus, this virus sequence does not meet the criteria to qualify as a new viral species. Nonetheless, these viral sequences have been deposited in GenBank due to the differences within the nucleotide sequences from published viral genomes and the identity of the host fungal strain. The RdRP and CP protein sequences from MgPV1, PcPV1, and TmPV1 do not share significant identity with known gammapartitivirus sequences and are isolated from new fungal hosts. As such, these three viruses are characterized as new viral species.



S5 Fig. Analyses of identified viral genomes belonging to the genus *Gammapartitivirus*. (A) Percent identity matrix, generated by Clustal-Omega 2.1, between identified viruses and select gammapartitivuses. The top half of the matrix, in the blue scale, is the percent identity of RdRP protein sequences and the bottom half of the matrix, in the green scale, is the percent identity of CP protein sequences. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the viruses identified. Where a sequences has a value greater than the species demarcation cutoff is noted, specifically >90% for RdRP and >80% for CP sequences. Bolded names are ICTV-recognized member species while italicized names are related, but unclassified viruses. Full names and accession numbers for protein sequences are in Table S2. (B) Density plot of read counts per nucleotide across the viral genome for the five gammapartitivuses. Graphical depiction of the genome structure of the RNA segments, predicted ORFs, and RdRP motif are on top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted for each virus. (C) Alignment of 5'UTR sequences of RNA1 and RNA2 for the identified viral sequences. Conserved nucleotides are highlighted with a blue box and number of nucleotides not shown in the alignment are noted in parentheses. Viruses identified in this study are noted with a red star. RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein.

New Genus: Epsilonpartitivirus

Two viral genomes, *Colletotrichum eremochloae partitivirus 1* (CePV1) (PRJNA262412; unpublished) and *Penicillium aurantiogriseum partiti-like virus* (PaPIV) (PRJNA369604; [41]), were identified that are phylogenetically related to an emerging group of partitiviruses that are separate from, but related to, the genus *Gammapartitivirus* (Fig 2); this group has been tentatively named *Epsilonpartitivirus* [41]. Identification of the PaPIV sequences was expected as this sample was RNA-seq data from *Cryphonectria parasitica* infected with PaPIV, previously isolated from *P. aurantiogriseum* by the same researchers [41]. To distinguish this viral genome from that isolated from the original host, the virus has been named PaPV1-cp. Also, the slight modification of the virus name, from partiti-like to partitivirus, demonstrates a recent improved understanding that this is a partitivirus following the identification of a RNA segment encoding a putative CP [41].

Members of this proposed genus have been identified from both fungal and invertebrate hosts, although currently the viral sequences form two sister clades separated by kingdom (Fig 2). Genome sizes of the dsRNA segments of putative epsilonpartitiviruses are from 1.7 to 1.9 kb for dsRNA1 and 1.5 to 1.8 kb for dsRNA2 (Table 1). Using a pairwise comparison analysis, a value of 85% identity for the RdRP protein sequence and 75% for the CP sequence would currently be appropriately stringent for species demarcation of viruses isolated from different fungal hosts (Fig 5A).

Observed here, sequences isolated from the PaPIV1-infected *C. parasitica* sequencing sample contain a number of SNPs in comparison to the source PaPIV sequence; these changes may reflect an adaptive response by the virus to a new fungal host [41]. A positive effect on fungal growth is observed whereby the virus confers a positive effect on fungal growth of *C. parasitica* in the presence of 2% salts that has not yet been confirmed in the original host *P. aurantiogriseum*. Read coverage along the viral genome segments isolated from *C. parasitica* is on average low, 18x for dsRNA1 and 5x for dsRNA1 (Fig 5B), similar to the published report [41].

Both of the viral genome segments of CePV1 share the greatest percent identity with PaPIV1, but below the threshold suggested above, at 84% and 73% for the RdRP and CP respectively (Fig 5A). Similar to other partitiviruses, a multiple sequence alignment of the 5'UTR sequences of the two genome segments of CePV1 reveals regions of sequence conservation (Fig 5C).

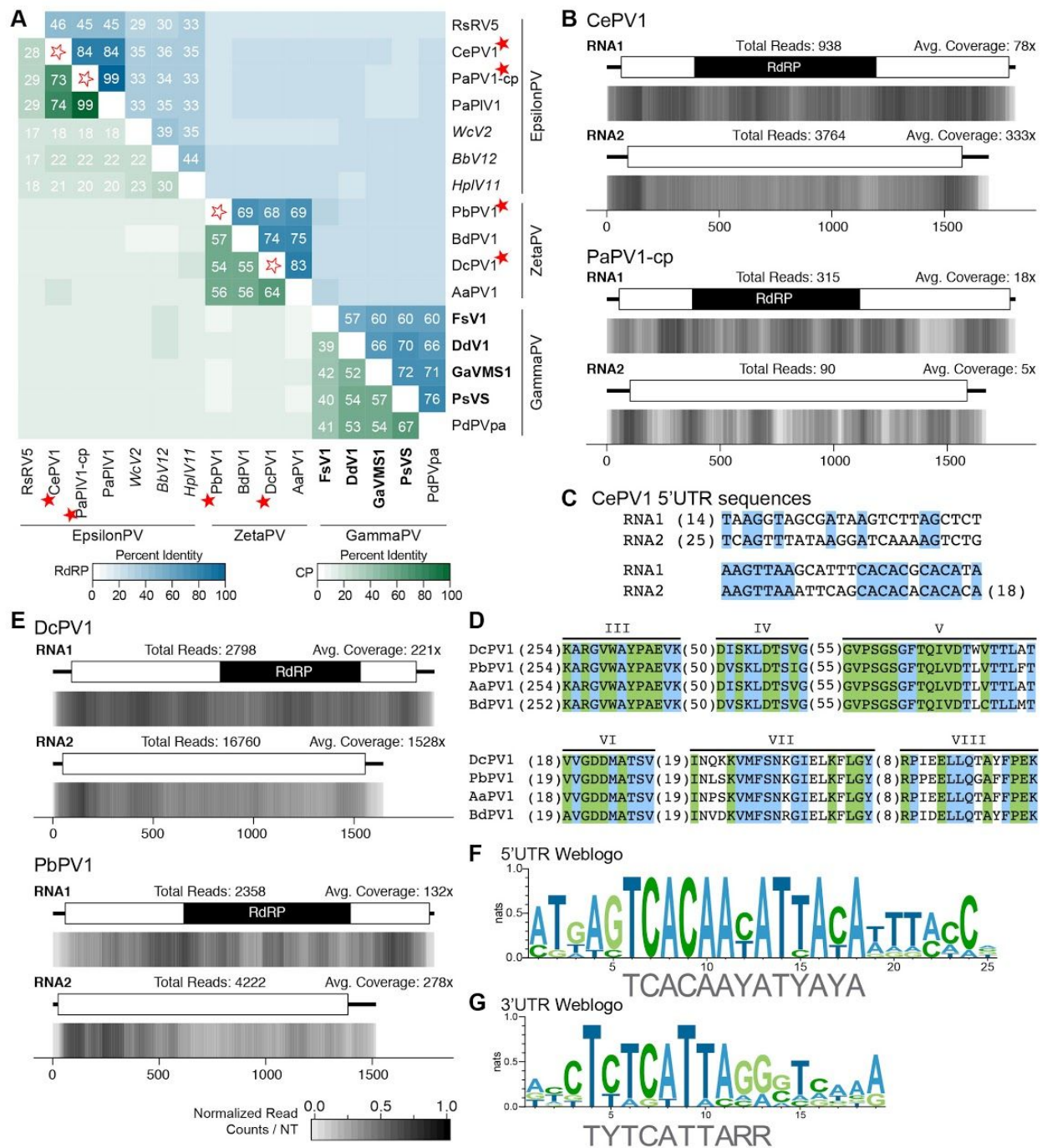


Fig 5. Analyses of identified viral genomes belonging to the two proposed genera, *Epsilonpartitivirus* and *Zetapartitivirus*, within the *Partitiviridae* family. (A) Percent identity matrix, generated by Clustal-Omega 2.1, between members of the two new groups and *gammapartitiviruses*, the most closely related, ICTV-recognized genus. The top half of the matrix, in the blue scale, is the percent identity of RdRP protein sequences and the bottom half of the matrix, in the green scale, is the percent identity of CP protein sequences. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the viruses identified. Two cutoff values were applied and as such only percent identities $\geq 29\%$ among RdRP sequences and $\geq 17\%$ among CP sequences are indicated. Virus names in italics refer to viruses isolated from insect, not fungal, hosts. Bolded virus names are member species recognized by ICTV. (B and E) Density plot of read counts per nucleotide across the viral genome for the two epsilonpartitiviruses (B) and two zetapartitiviruses (E). Graphical depiction of the genome structure, predicted ORFs, and RdRP motif are along the top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1; heatmap scale found in (E) applies also to (B). Total reads mapped to the virus genome and average read coverage are noted for each virus. (C) Alignment of the 5' UTR sequences from the two genomic segments of the epsilonpartitivirus CePV1. Conserved nucleotides are highlighted in blue and numbers in parentheses indicate nucleotides not shown. (D) Alignment of the six conserved motifs within the RdRP protein sequence of all four proposed zetapartitiviruses. Amino acids highlighted in blue are only common among zetapartitiviruses while residues in green are shared with gammapartitiviruses. (F and G) Weblogos generated from alignments of the (F) 5'UTR and (G) 3'UTR sequences of both RNA genome segments for all four viruses in the Zetapartitivirus genus. Red stars indicate viruses identified in this study. RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein.

Initially a mycoviral sequence was isolated from a *Colletotrichum navitas* sequencing sample (PRJNA262223; unpublished); however, other than small variances within the 5' and 3' UTRs, the nucleotide sequence is identical to that of CePV1. It is possible that two different, closely related, fungi are infected with the same virus, however as noted above with respect to the PaPIV1-infected *C. parasitica*, SNPs can quickly appear within a viral genome sequence when in a new host. Therefore further analyses were performed to determine the likelihood of a CePV1-like sequence hosted by *C. navitas*. First it was determined that, after controlling for the unaligned reads pool size, 2.6x more viral reads were present in the *C. eremochloae* sample than the *C. navitas* sample. Second, to identify possible cross-contamination between the two species within the sequencing data, non-virus-derived reads were reciprocally aligned to the other fungal genome. Thus reads from *C. navitas* were aligned to the *C. eremochloae* genome, and vice versa. Here we observed that 5% of reads from *C. eremochloae* aligned to the *C. navitas* genome while 65% of reads in the *C. navitas* pool aligned to the *C. eremochloae* genome. Evaluating both of these results, we propose that the likely host of this viral sequence is *C. eremochloae* and the identification of the same sequence within the *C. navitas* sample is likely due to cross contamination from the *C. eremochloae* sample. Despite the possible ambiguity in fungal host for this viral sequence, we recommend its inclusion in the recently suggested genus *Epsilonpartitivirus* as a new viral species.

New Genus: Zetapartitivirus

The final two viruses identified within the *Partitiviridae* branch of the phylogenetic tree (Fig 2) are *Penicillium brasilianum partitivirus 1* (PbPV1) (PRJEB7514; [42]) and *Delitschia confertaspora partitivirus 1* (DcPV1) (PRJNA250740; unpublished). Both viruses, the first to be described in these fungal species, phylogenetically group with *Alternaria alternata partitivirus 1* (AaPV1) [43] and *Botryosphaeria dothidea partitivirus 1* (BdPV1) [44], a distinct clade within the *Partitiviridae* that is related to, but separate from, the genus *Gammapartitivirus* (Fig 2).

These four viruses share many sequence-based similarities. A 13-nt highly conserved sequence (TCACAAATYAYA) is present in the 5'UTR sequence of both dsRNA genome segments from all four viruses (Fig 5F). Additionally, a 10-nt conserved sequence (TYTCATTARR) is found within the 3'UTR sequence of both dsRNA segments of all four viruses (Fig 5G). While these four viruses contain the six conserved motifs common to RdRP proteins from other members of the *Partitiviridae*, including the GDD triad in motif VI, there are distinct differences that are shared that distinguish these viruses from the existing genera (Fig 5D).

Further, the range of lengths for the RdRP protein, from 569 to 572 aa, and CP protein, from 453 to 501 aa (Table 1), are unique ranges among the currently described genera [30]. A pairwise percent identity analysis of the RdRP and CP sequences from the four viruses suggests a threshold for species demarcation at 90% for the RdRP sequence and 80% for the CP sequence, similar to other genera within the family (Fig 5A).

Alignment of the RNA-seq reads to the viral genomes demonstrates that dsRNA2 has a higher average coverage per nucleotide for both viruses, with approximately 2x more coverage observed between the segments of PbPV1 and 7x between the segments of DcPV1 (Fig 5E).

We recommend that DcPV1 and PbPV1, along with AaPV1 and BdPV1 belong to a new genus within the family *Partitivirus*, called *Zetapartitivirus*. The genomic segment and protein sequence lengths are consistent and unique within the *Partitivirus* family, and the RdRP protein sequences form a distinct clade that is separate from the closest genus, *Gammapartitivirus* (Fig 2).

Chrysoviridae

To date, nine species have been classified as belonging to the Chrysoviridae family of dsRNA viruses that infect either ascomycota or basidiomycota fungi [45]. Members of this family have multipartite genomes, consisting of three to five segments. Here we identify two viral genomes with sequence similarity to Chrysoviruses: *Beauveria bassiana chrysovirus 1* (BbCV1) ([46]PRJNA306902) and *Penicillium raistrickii chrysovirus 1* (PrCV1) (unpublished; PRJNA250734) (Fig 2).

Phylogenetically, chrysoviruses are grouped into two distinct clusters, where cluster 1 contains members with either three or four segmented genomes, while cluster 2 contains viruses with either four or five segments [45]. The two viral sequences identified in this study, BbCV1 and PrCV1, belong with cluster 1 chrysoviruses; both genomes are comprised of four genomic segments and the RdRP sequences phylogenetically group with the ICTV-recognized members within cluster 1, *Isaria javanica chrysovirus 1* (IjCV1) [47] and *Penicillium chrysogenum virus* (PcCV1) [48] (Fig 2). Further, as detailed below, we recommend that these are both new viral species within the genus *Chrysovirus*.

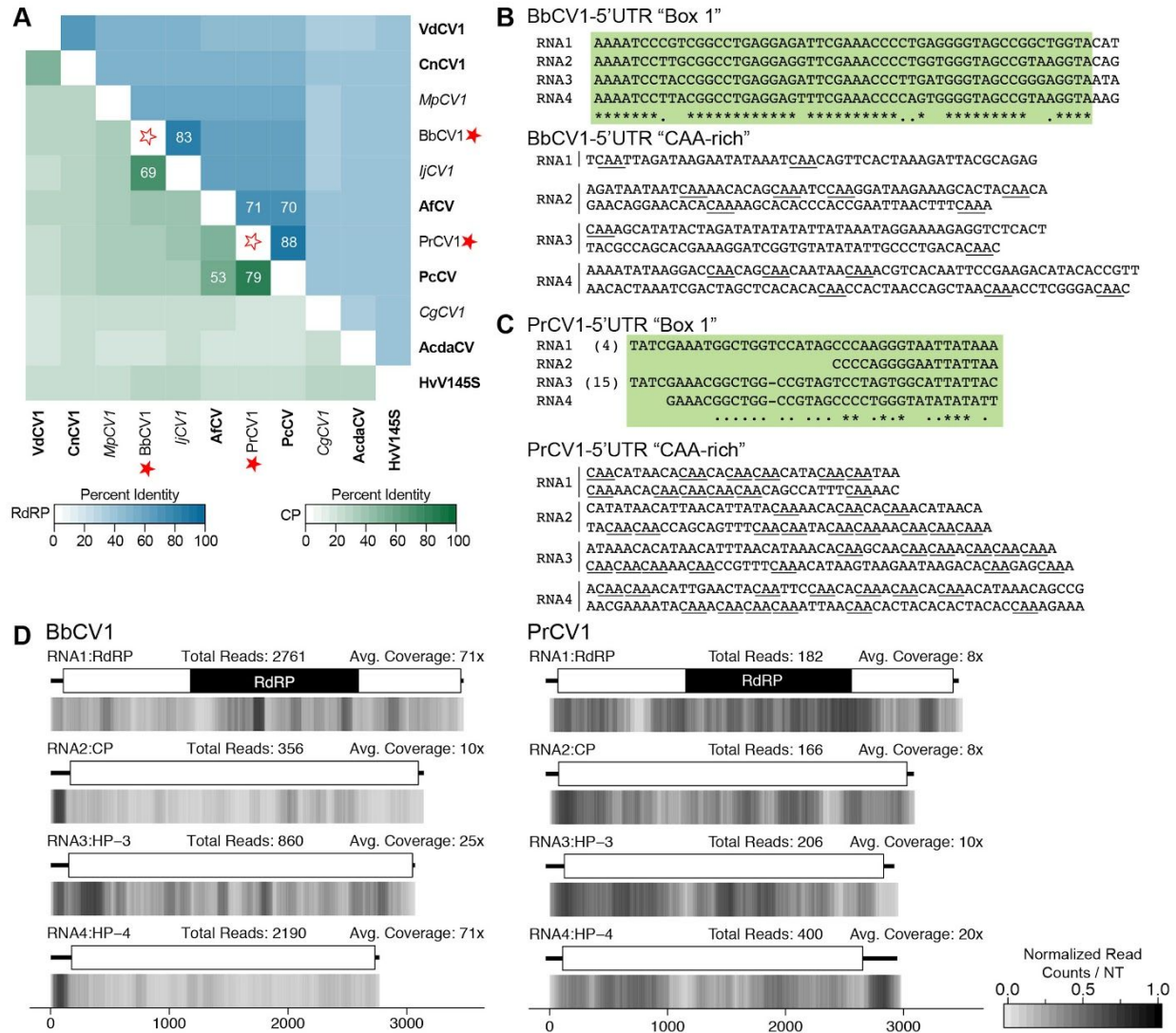
To our knowledge, these are the first chrysoviruses identified for both of these fungal hosts, and indeed the first virus from *P. raistrickii* described, thus satisfying the host-based criteria for species demarcation. Following identification of the RdRP-encoding contig, the identity of the other three segments was determined via a BLASTP search using protein sequences of known

chrysovirus against the predicted protein sequences of the Trinity contigs. The combined length of dsRNAs for both BbCV1 and PrCV1 are 12.46 and 12.56 kb respectively (Table 1), within the observed range of other 11.5 to 12.8 kb of other chrysovirus [45]. In general, chrysovirus genome segments are numbered by size, however segment 4 of PrCV1 is slightly longer than segment 3 (Table 1), similar to segment 4 of *Amasya cherry disease associated chrysovirus* (AcdaCV) which is also larger than its respective segment 3 [49]. A BLASTP analysis confirms that segment 4 of PrCV1 best matches the segment 4 of *Penicillium chrysogenum virus* and thus is numbered accordingly. An alignment of the RNA-seq reads to the viral genome segments reveals uniform average coverage per nucleotide for the four genomic segment of PrCV1 (from 15x to 32x) and slightly more variable average coverage of BbCV1 segments, ranging from 10x to 72x (S6D Fig). The dsRNA genomic segment lengths therefore also satisfy the criteria for species demarcation within the genus *Chrysovirus*.

The 5'UTR of chrysovirus is generally between 140 and 400 nt in length and contains two regions of high sequence conservation, Box 1 and Box 2 [45], although a number of recently characterized viruses have shorter 5'UTRs, including *Cryphonectria nitschkei chrysovirus 1* (CnCV1) which has 82 nt 5'UTR for segment 1 [50] and AcdaCV which has 5'UTR lengths of 87, 96, 95, and 106 bp [49]. With the exception of the 5'UTR of BbCV1-segment 1 (106 nt), the 5'UTR of all segments discovered for BbCV1 and PrCV1 fall within the published range (S6B Fig and S6C Fig). A highly conserved region is present within area corresponding to "Box 1" within the 5'UTR of the four genome segments of BbCV1 (S6B Fig, top panel) and the expected "CAA" sequence is also observed in all sequences (S6B Fig, bottom panel). Similarly both conserved regions are observed within the 5'UTR of PrCV1 (S6C Fig). Thus the size and composition of the 5'UTR sequences of BbCV1 and PrCV1 also satisfy the criteria for chrysovirus classification.

The final criteria considered for species demarcation is the percent identity of the RdRP and CP protein sequences relative to other members within the genus, where the thresholds are 70% and 53% respectively [45]. BbCV1 RdRP and CP share 83% and 69% identity with the related proteins of IjCV1, while PrCV1 RdRP and CP protein sequences share 88% and 79% identity with those of *Penicillium chrysogenum virus* (PcCV1) (S6A Fig). Since BbCV1 was isolated from a new fungal host, we propose classifying it as a new species of chrysovirus, while the relative relatedness between the hosts *P. raistrickii* and *P. chrysogenum* suggest that PrCV1 is not a new species of chrysovirus. Further identification of additional viruses plus biological

characterization of host specificity will provide the needed insights to properly classify viruses within this family.



S6 Fig. Analyses of viral genomes belonging to the *Chrysoviridae* family. (A) Percent identity matrix, generated by Clustal-Omega 2.1, versus select members of the *Chrysovirus* genus. The top half of the matrix, in the blue scale, is the percent identity of RdRP protein sequences and the bottom half of the matrix, in the green scale, is the percent identity of CP protein sequences. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the viruses identified. Values above 70% for RdRP sequences and 53% for CP sequences are indicated. ICTV-member and unclassified viruses are noted in bold and italics respectively. (B and C) Alignment of the 5'UTR sequences for the four genomic segments of BbCV1 (B) and PrCV1 (C). Two highly conserved regions are found in the 5'UTR; the first is a 40-75 nt region, highlighted in green, while the second region is rich in CAA repeats (underlined). (D) Density plot of read counts per nucleotide across the viral genome segments. Graphical depiction of the genome structure and predicted ORFs and motifs are on top and on the bottom is a heatmap of read counts normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. Red stars indicate viruses identified in this study. RdRP: RNA-dependent RNA Polymerase; CP: Coat Protein; HP: Hypothetical Protein.

Endornaviridae

Members of the *Endornaviridae* family of viruses have long, single segment, dsRNA genomes, generally encoding a single polyprotein. In addition to the RdRP domain and helicase domain (Viral helicase 1 or DEADx helicase) found in all endornaviruses, viral genomes may also encode a viral methyltransferase domain and/or a glycosyltransferase domain. Recently a two-genus split was proposed [51] where characteristics of clade I (*Alphaendornavirus*) include a plant-host, larger genomes (ranging from 13.8 to 17.5 kb), the presence of a UGT domain and a site-specific nick towards the 5' end of the genome sequence [52], while members of clade II (*Betaendornavirus*) have fungal-hosts and generally have genome sizes ranging from 9.5 to 11.4 kb, encode a viral methyltransferase domain but not a UGT domain, and lack the site-specific nick [51].

Here four new virus sequences were identified that group with known endornaviruses (Fig 6a), *Gyromitra esculenta endornavirus 1* (GeEV1) (PRJNA372840; unpublished) and *Morchella*

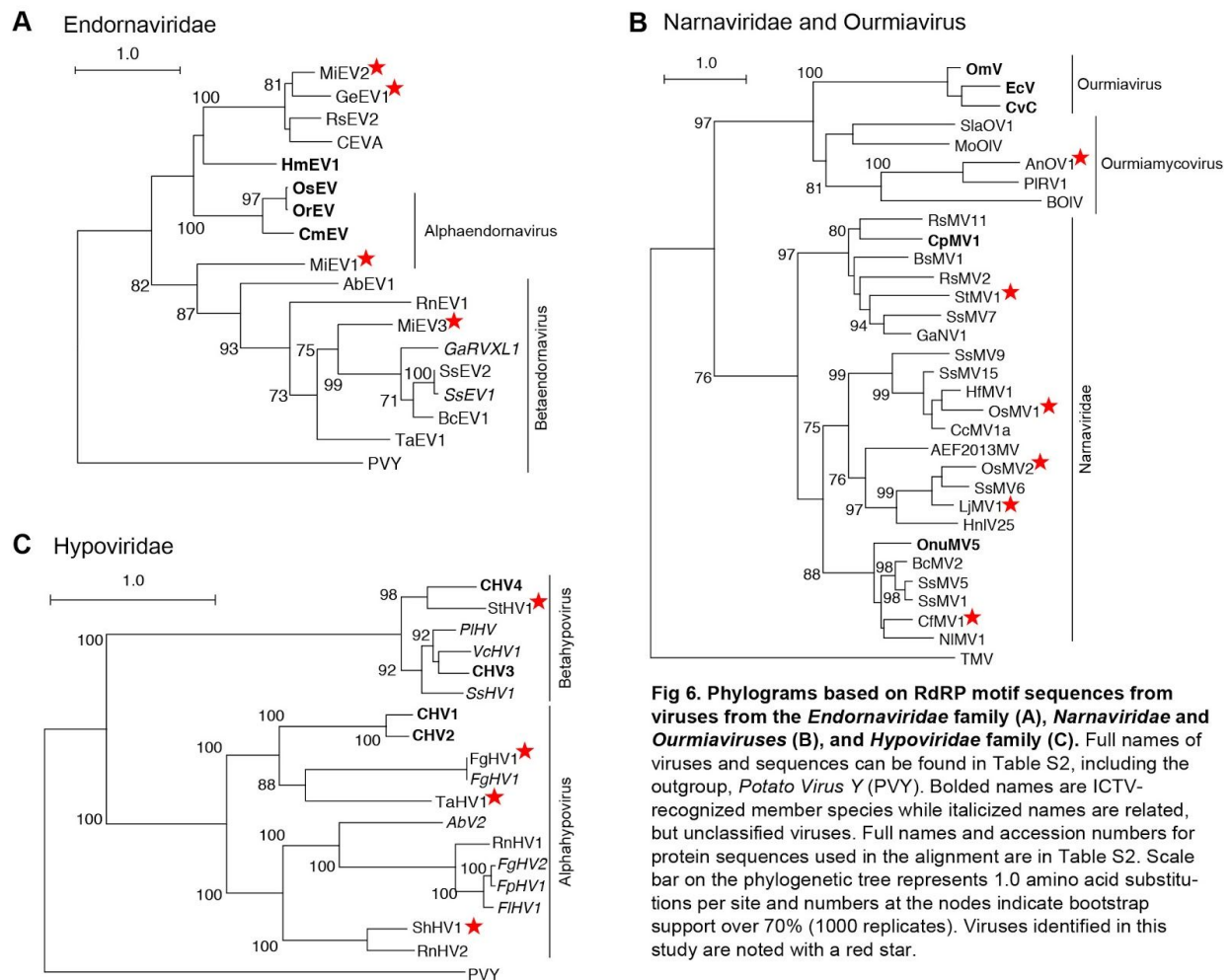


Fig 6. Phylograms based on RdRP motif sequences from viruses from the Endornaviridae family (A), Narnaviridae and Ourmiaviruses (B), and Hypoviridae family (C). Full names of viruses and sequences can be found in Table S2, including the outgroup, *Potato Virus Y* (PVY). Bolded names are ICTV-recognized member species while italicized names are related, but unclassified viruses. Full names and accession numbers for protein sequences used in the alignment are in Table S2. Scale bar on the phylogenetic tree represents 1.0 amino acid substitutions per site and numbers at the nodes indicate bootstrap support over 70% (1000 replicates). Viruses identified in this study are noted with a red star.

importuna endornavirus 1, 2, and 3 (MiEV1, MiEV2, MiEV3) (PRJNA372858; unpublished).

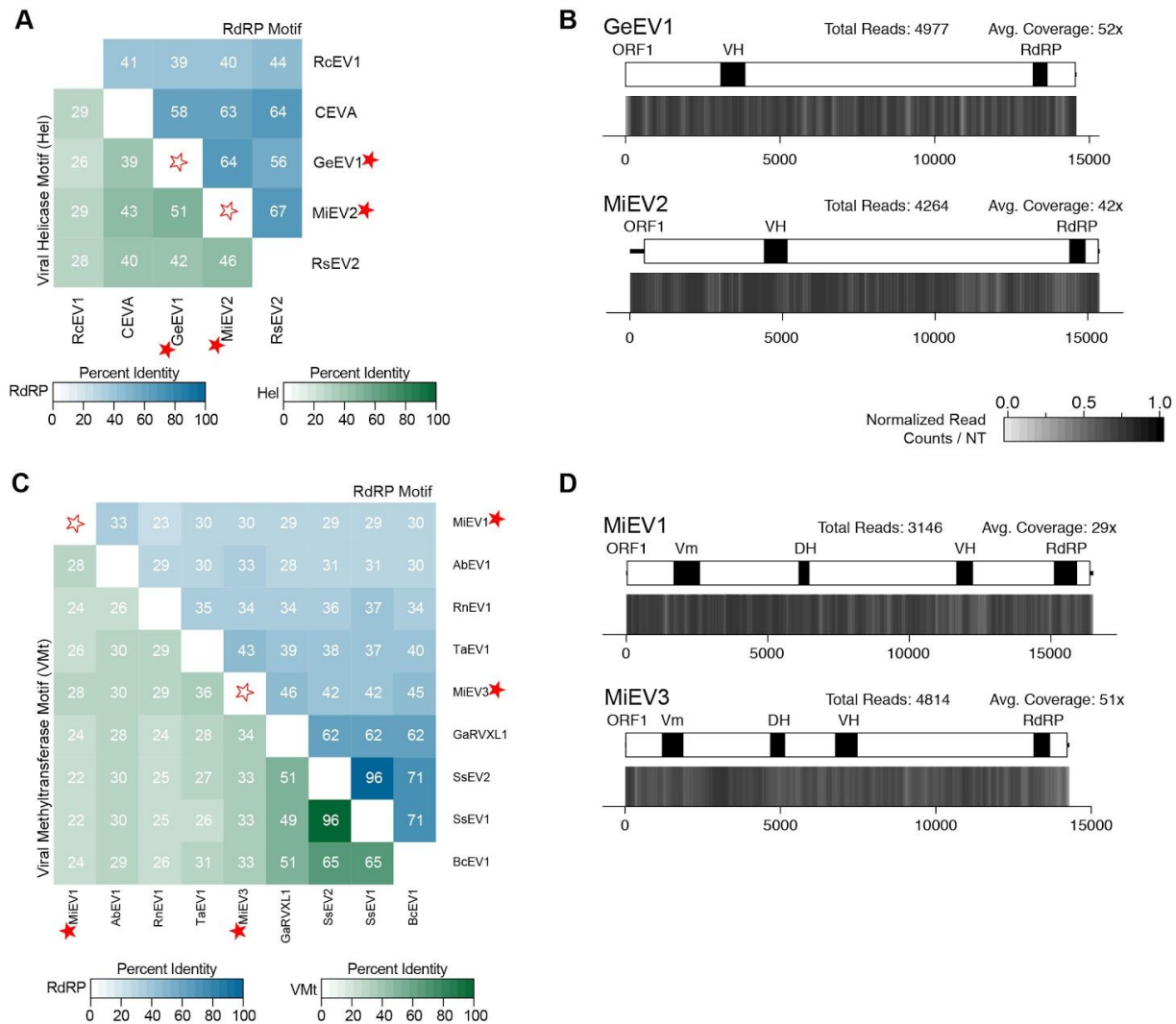
These are the first viruses, to our knowledge, identified from both *G. esculenta* and the morel mushroom, *M. importuna*.

Alphaendornavirus

A phylogenetic analysis of the RdRP motif from related endornaviruses and the viruses identified in this study reveals MiEV2 and GeEV1 form a well supported clade with other fungal endornaviruses, including *Ceratobasidium endornavirus A* (CEVA) [53], and *Rhizoctonia solani endornavirus 2* (RsEV2) [54], that is related to, but separate from alphaendornaviruses isolated from plant hosts, including the type species *Oryza sativa endornavirus* (OsEV) [55] (Fig 6A). The genome sizes of MiEV2 and GeEV1 are 15.3 and 14.5 kb respectively, well within the observed range of other *Alphaendornaviruses* (Table 1). Additionally, a conserved domain search within the protein coding sequences determined both viruses encode a viral helicase and an RdRP domain, but not a viral methyltransferase domain, nor a UGT domain (S7B Fig). This combination of protein domains appears to be shared among mycoviruses and not plant-hosted endornaviruses.

GeEV1 has a predicted AUG at position 5, resulting in a shorter 5' UTR sequence than observed in other endornaviruses. The predicted start codon for MiEV2 however is at position 476, resulting in an unexpectedly long 5'UTR; as the entire sequence upstream of this codon is in-frame with the coding sequence, it is possible that this genome sequence is partially truncated and is missing the true AUG sequence. As manual analysis of the RNA-seq data did not result in an extension of the 5' region of MiEV2, we recommend further analysis of the biological material to confirm the 5' region of this virus. An alignment of RNA-seq reads to the two viral genomes reveals uniform coverage across both genome sequences, with a total of 4977 reads and 8171 reads respectively (S7B Fig).

The two criteria for species demarcation within the genus *Alphaendornavirus* are difference in host and an overall nucleotide sequence identity below 75%. Both viruses described here satisfy these criteria. Indeed, nucleotide identity among *Endornaviruses* is generally below 50%. An analysis of the protein sequences of the RdRP motif and the viral helicase motif reveal similar results, as sequences are between 39% and 67% identical for the RdRP sequences and 26% to 51% identical among the helicase sequences (S7A Fig). Together these analyses support the recognition of GeEV1 and MiEV2 as new species within the *Endornaviridae* family, and the genus *Alphaendornavirus* in particular.



S7 Fig. Analyses of identified viral genomes belonging to the family Endornaviridae. (A and C) Percent identity matrix, generated by Clustal-Omega 2.1, of mycoviruses in the genus *Alphaendornavirus* (A) and *Betaendornavirus* (C). The top half of both matrices, in the blue scale, is the percent identity of the RdRP motif sequences; the bottom half of the matrix, in the green scale, is the percent identity of the Viral Helicase 1 motif (A) and Viral methyltransferase motif (C). For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the viruses identified in this study. (B and D) Density plot of read counts per nucleotide across the viral genome for the alphaendornaviruses, GeEV1 and MiEV2 (B) and betaendornaviruses MiEV1 and MiEV3 (D). Graphical depiction of the genome structure, predicted ORF, and encoded motifs are on top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. Viruses identified in this study are noted with a red star. RdRP: RNA-dependent RNA Polymerase; VH: Viral Helicase 1; Vm: Viral methyltransferase; DH: DEXDx Helicase.

Betaendornavirus

The two other endornaviruses identified were MiEV1 and MiEV3, which group with the betaendornaviruses. MiEV3, in particular, forms a strong phylogenetic relationship with other characterized betaendornaviruses, while MiEV1 is the outermost member of this clade (Fig 6A). Unexpectedly, the genome sizes of both viruses are more than 25% larger than other

betaendornaviruses; MiEV1 has a genome size of 16.5 kb while MiEV3 has a genome size of 14.3 kb (Table 1). However, the motifs within the protein sequences are as expected, as both MiEV1 and MiEV3 contain a viral methyltransferase domain, helicase domain(s), and an RdRP domain (S7D Fig). Specifically MiEV1 and MiEV3 both encode a DEADx helicase domain and a viral helicase 1 domain, similar to the type species *Sclerotinia sclerotiorum endornavirus-1* (SsEV1) (NC_021706.1). Alignment of the RNA-seq reads to the viral genomes reveals an average coverage per nucleotide of 29x and 51x for MiEV1 and MiEV3 respectively (S7D Fig), and coverage appears uniform across both genomes.

Similar to the alphaendornaviruses, nucleotide identity is generally below 50%, with the exception of SsEV1 and *Sclerotinia sclerotiorum endornavirus 2* strain DIL-1 (SsEV2) which share 82% nucleotide identity. Further analysis of the protein sequences from the RdRP motifs and viral methyltransferase motifs determined that a percent identity threshold for either motif of 75% would be appropriate for species demarcation (S7C Fig). In particular, both MiEV1 and MiEV3 are well below both a nucleotide-based or a protein-motif based threshold. Thus, despite the same fungal host, the uniqueness of the viral genomes sequence for both viruses suggests that both MiEV1 and MiEV3 are new species of the family *Endornaviridae*, and belong within the genus *Betaendornavirus*.

Positive(+)ssRNA Viruses

Tobamo-like virus: ArTIV1

A virus sequence was identified in an *Acidomyces richmondensis* sample (PRJNA250470; unpublished) with a single segment genome sequence, 10,291 nt in length, that encoded four predicted ORFs (Table 1). The gene product from ORF2 contains the RdRP domain (Fig 7B) and a BLAST-P search of the 'nr' database with this sequence identified two related, full length mycoviruses, *Macrophomina phaseolina tobamo-like virus 1* (MpTIV1) [5] and *Podosphaera prunicola tobamo-like virus 1* (PpTIV1) [56]. As such, this newly identified virus has been named *Acidomyces richmondensis tobamo-like virus 1* (ArTIV1).

For all three mycoviral sequences the first and largest ORF contains a viral methyltransferase and viral helicase domain and ORF2 encodes the RdRP domain (Fig 7B). ORF3 and ORF4 in MpTIV1 have been putatively identified as encoding a movement protein and coat protein respectively [5].

A phylogenetic analysis of the RdRP motif from the three mycoviruses compared to select members of the families *Virgaviridae*, *Bromoviridae* and genus *Idaeovirus* demonstrated that ArTIV1 is more closely related to the two previously identified mycoviruses than known plant viruses (Fig 7A). Further, the group of mycoviruses form a sister clade to the *Virgaviridae* viruses. Of the viruses currently characterized from this family, only members of the genus *Tobamovirus* have a non-segmented genome similar to ArTIV1 and relatives; other members have multipartite genomes [57].

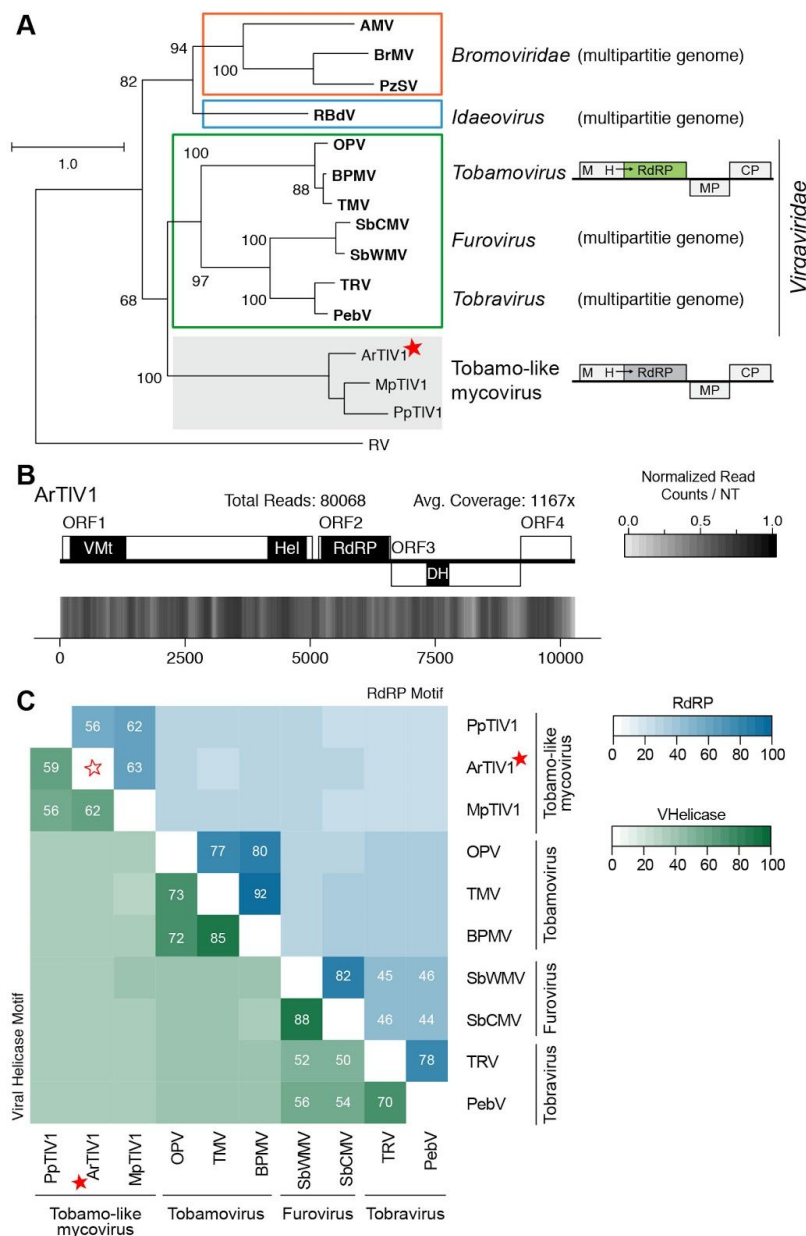


Fig 7. Analyses of *Acidomyces richmondensis tobamo-like virus 1* (ArTIV1) and related viruses. (A) Phylogram of the RdRP-motif for ArTIV1, related mycoviruses, and select members of the genus *Tobamovirus* and family *Bromoviridae*; sequence from *Rubella Virus* (RV) serves as the outgroup. Bolded names are ICTV-recognized member species. Full names and accession numbers for protein sequences used are in Table S2. Numbers at the nodes indicate bootstrap support values $\geq 65\%$ (1000 replicates). Scale bar on the phylogenetic tree represents 1.0 amino acid substitutions per site. Diagrammatic genome structures are to the right of the phylogram. (B) Density plot of read counts per nucleotide across the ArTIV1 viral genome. Graphical depiction of the genome structure, predicted ORFs, and encoded motifs are along the top and a heatmap of normalized read counts is on the bottom. Reads are normalized to a 0 - 1 scale, where the maximum read count equals 1. Total reads mapped to the virus genome and average read coverage are noted. (C) Percent identity matrix, generated by Clustal-Omega 2.1. The top half of the matrix, in the blue scale, is the percent identity of the RdRP motif sequence while the bottom half of the matrix, in the green scale, is the percent identity of the residues from Viral Helicase motif. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of ArTIV1. Cutoff values of 44% and 40% were applied to the RdRP and Helicase matrices respectively. Viruses identified in this study are noted with a red star. VMt: Viral Methyltransferase; VHel: Viral Helicase; RdRP: RNA-dependent RNA Polymerase; DH: DEXDc Helicase; ORF: Open Reading Frame.

To further characterize this emerging group of mycoviruses, a percent identity analysis was undertaken, comparing the amino acid sequences from the RdRP and viral helicase motifs to members of the *Virgaviridae* family. The RdRP motif sequences among the mycoviruses share 56 to 63% identity with one another, but less than 44% identity with *Virgaviridae* viruses (Fig 7C). Further, sequences from the viral helicase domain were 59 to 62% identical among mycoviruses, and less than 40% shared identity outside of this clade. Due to these values, it appears that all three viruses are unique viral sequences.

Finally, the RNA-seq data was aligned to the ArTIV1 genome sequence, which revealed uniform coverage along the length of the genome, averaging 1,167 reads per nucleotide (Fig 7B).

To date all viruses of the *Virgaviridae* family infect a plant host, although many members are transmitted by nematodes or plasmodiophorids, a group of parasitic microscopic organisms unrelated to fungi or oomycetes [57]. Tobacco mosaic virus (TMV), the best characterized *Tobamovirus*, can infect, replicate, and persist in the ascomycota fungi *Colletotrichum acutatum* [58]. Further, a virus from the *Bromoviridae* family, Cucumber mosaic virus (CMV) was shown to naturally infect the phytopathogenic fungi *Rhizoctonia solani* and as well as replicate within *Valsa mali* but not *C. parasitica*, and *F. graminearum* [59]. Due to the close contact between phytopathogenic fungi and plants a change in host range of a tobamo-like virus is plausible. Interestingly, the *A. richmondensis* strain analyzed in this study was isolated from acid mine drainage biofilms [60] and is not known to be phytopathogenic. Acid mine drainage biofilms are complex communities comprised of different fungi along with bacteria, archaea and bacteriophages [61,62]. The role of the mycovirus identified in this study in such an extreme and heterogeneous environment remains to be determined.

Ambiguiviridae

Four viral genome sequences were identified that share similarity to unclassified (+)ssRNA mycoviruses: *Trichoderma harzianum ssRNA virus 1* (ThAV1) (PRJNA216008; [63]), *Setosphaeria turcica ssRNA virus 1* (StAV1) (PRJNA250530; unpublished), *Verticillium longisporum ssRNA virus 1* (VIAV1) (PRJNA308558; unpublished), and *Periconia macrospinoso ssRNA virus 1* (PmAV1) (PRJNA262386; [64]). To our knowledge, this is the first viral sequence described for the later three fungal species, and the first ssRNA virus isolated from *T. harzianum*.

Six other viruses appear related to these four newly identified viruses: *Diaporthe ambigua* RNA virus 1 (DaRV1), *Magnaporthe oryzae* RNA virus (MoRV), *Soybean leaf-associated ssRNA virus 1 and 2* (SlaRV1, SlaRV2), *Verticillium dahliae* RNA virus (VdRV), and *Sclerotinia sclerotiorum umbra-like virus 1* (SsUIV1) [5,6,65–67]. The RdRP protein sequence of this group of mycoviruses shares some sequence similarity to RdRPs of recently described (+)ssRNA viruses from insects, as well as members of the plant-infecting family *Tombusviridae*. Phylogenetic analysis demonstrates the fungal viruses form a well-supported clade that is separate from both the insect and plant viruses (Fig 8A). A pairwise identity analysis of the RdRP motif region from the viruses analyzed in the phylogenetic tree reveals a range of identities between the fungal viruses, from 36% to 63%, while only sharing 25% to 33% identity with the insect viruses and 26% to 32% identity with the members of the *Tombusviridae* (Fig 8D). A similar analysis using the ORF1 protein coding sequence, from only the mycoviral genomes, demonstrates 18 to 38% identity among sequences (Fig 8D).

Together these ten fungal viruses are characterized by a single segment genome, 2.6 kb to 4.5 kb in length, that encodes two ORFs, where the second ORF contains the RdRP domain (Table 1). The amber stop codon 'UAG', found at the end of the first ORF, has been suggested to result in readthrough to produce a fusion protein with the downstream RdRP [67]. All members of this group of viruses contain the amber stop codon, and both ORF1 and ORF2 are found in the same frame, a requirement for readthrough. Additionally, all 10 viruses share a number of key differences in residues commonly conserved within (+)ssRNA viruses. For instance, the second aspartic acid residue in motif A, a feature of (+)ssRNA viruses [68], is instead a glutamic acid (Fig 8C). Additionally, these 10 viruses have a GDN triad in motif C versus the canonical GDD found in (+)ssRNA viruses (Fig 8C). This triad has previously been observed within some (-)ssRNA viruses [68], although within (+)ssRNA viruses, modification of the GDD to GDN has repeatedly demonstrated a detrimental effect on enzymatic activity [69–71]. Finally, within motif D, a positively-charged lysine, common to nearly all RdRP sequences [68,72], is a nonpolar proline or valine residue (Fig 8C).

The 5'UTRs range in size from 283 bp to 889 bp in length for all the mycoviruses in this group, with the exception of VdRV1 and VIAV1, the two viruses isolated from species of *Verticillium*; these 5'UTRs are 11 nt and 54 nt respectively. Similarly, the 3'UTR sequences range in length from 372 bp to 1279 bp, while those of VdRV1 and VIAV1 are 182 nt and 29 nt respectively (Fig 8C). Average coverage per nucleotide across the viral genomes of the four

identified viruses is fairly uniform across the genome with ranges from 15x for VIAV1 to 734x for PmAV1 (Fig 8B).

The characteristics of these four identified viruses and their closest mycoviral relatives supports the creation of a new family of viruses within the (+)ssRNA class. We propose the

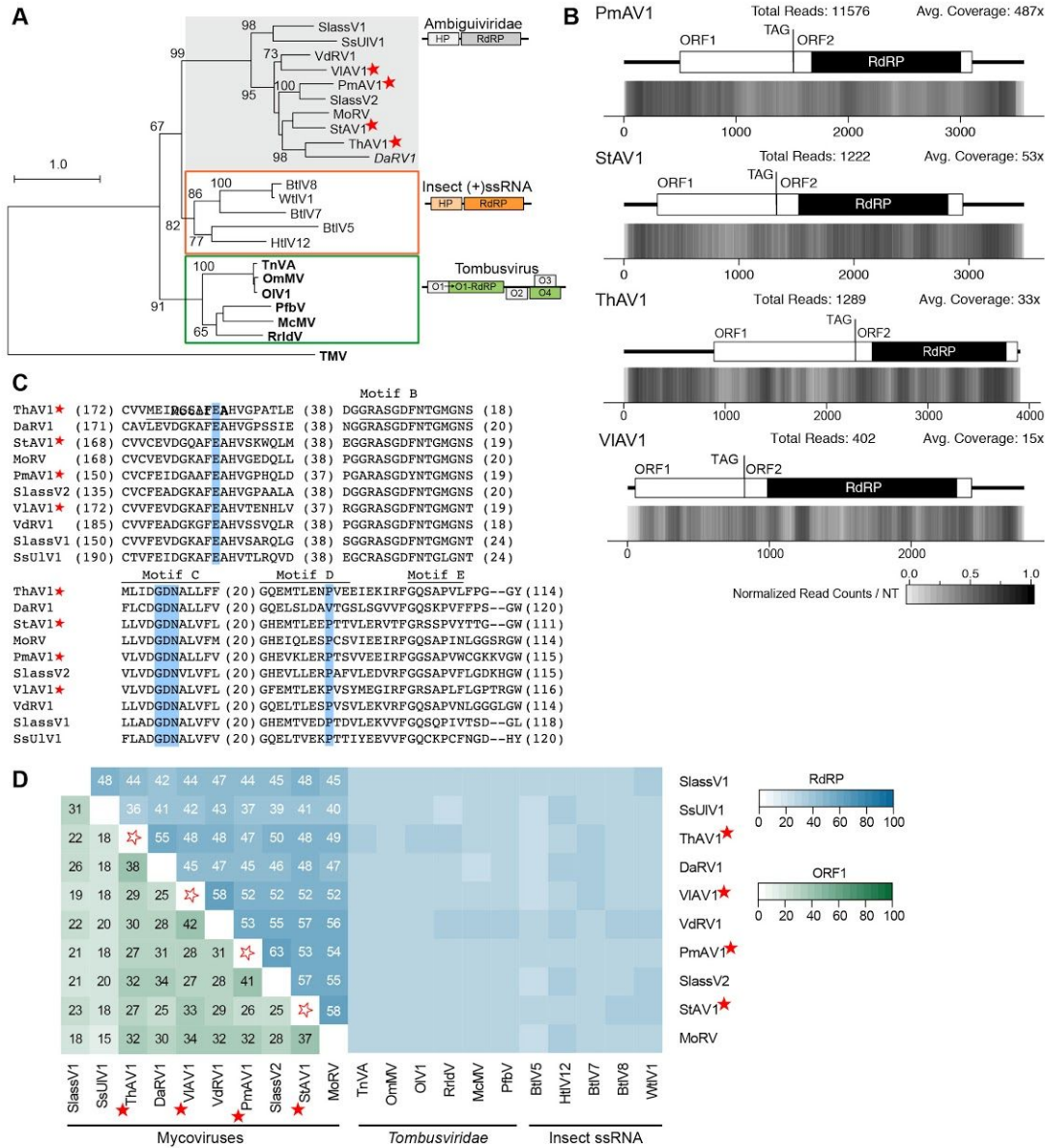


Fig 8. Analyses of four newly identified (+)ssRNA viral genomes and viruses belonging to the proposed family *Ambiguiviridae*. (A) Phylogram of the RdRP-motif for four new viruses and related ambiguiviruses along with select members of the family *Tombusviridae* and members of an unnamed group of (+)ssRNA insect viruses; sequence from *Tobacco Mosaic Virus* (TMV) serves as the outgroup. Bolded names are ICTV-recognized member species while italicized names are recognized, but unclassified viruses. Full names and accession numbers for protein sequences used are in Table S2. Numbers at the nodes indicate bootstrap support values $\geq 60\%$ (1000 replicates). Scale bar on the phylogenetic tree represents 1.0 amino acid substitutions per site. Diagrammatic genome structures and predicted ORFs for each clade of viruses is to the right of the phylogram. (B) Density plot of read counts per nucleotide across four viral genomes. Graphical depiction of the genome structure, predicted ORFs, encoded motifs and the amber stop codon (TAG) for each genome are along the top and a heatmap of normalized read counts is on the bottom. Reads are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. (C) Multiple sequence alignment of the RdRP motifs of members of the *Ambiguiviridae* family. Motifs A-E found in (+)ssRNA viral RdRP sequences are indicated along the top of the alignment. Conserved residues that differ from canonical (+)ssRNA RdRP sequences, including the unusual GDN triad found in Motif C, are highlighted in blue. Numbers in parentheses are residues not shown. (D) Percent identity matrix, generated by Clustal-Omega 2.1, of mycoviruses that comprise the proposed family *Ambiguiviridae*. The top half of the matrix, in the blue scale, is the percent identity of the RdRP motif sequence and the bottom half of the matrix, in the green scale, is the percent identity of the ORF1 amino acid sequence. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note junction point for the row and column of the viruses identified. All RdRP motif sequences share $\geq 35\%$ identity and ORF1 protein sequences share $\geq 15\%$ identity. Viruses identified in this study are noted with a red star. RdRP: RNA-dependent RNA Polymerase; ORF: Open Reading Frame; HP: Hypothetical Protein.

establishment of the family *Ambiguiviridae* that contains DaRV1 and other recently discovered viruses related to it, including the four described in this study. The proposed name references the host of the first virus identified (*D. ambigua*), while also highlighting the unusual nature of the GDN triad within the RdRP motif.

Yadokariviridae

Samples *Aspergillus homomorphus* (PRJNA250984; unpublished) and *Penicillium digitatum* (PRJNA352307; unpublished), which yielded a totivirus (AhoTV1) and partitivirus (PdPV1) respectively, each also contained a second, unrelated contig containing an RdRP motif. BLAST-X analysis of these nucleotide sequences determined they were both related to a group of unclassified, (+)ssRNA viruses, that includes four other mycoviruses. As described below, these two viruses are new members of a recently proposed family, *Yadokariviridae* [73].

Aspergillus homomorphus yadokarivirus 1 (AhoYV1) and *Penicillium digitatum yadokarivirus 1* (PdYV1), along with other members of this group, are characterized by having single segment genomes that encode a single predicted ORF, the RdRP (Table 1). Genome sizes range from 3,144 nt (PdYV1) to 5,089 (*Yado-kari virus 1*, YkV1). The length of the RdRP protein sequence is predicted to range from 962 aa for *Aspergillus foetidus slow virus 2* (AfSV2) [74] to 1,430 aa for YkV1 [75]. The lengths of the 5'UTRs are all greater than 500 nt, except for PdYV1. Attempts to extend the sequence of the 5'UTR of PdYV1 using the RNA-seq data were not successful; however, as the average coverage per nucleotide for PdYV1 was low, just 23x (Fig 9C), additional experiments with the fungal source material may be fruitful. Lengths of the 3'UTRs for all six viruses are also largely similar, between 200 to 400 nt, except for YkV1 which has a 3'UTR of 1,223 nt. A genome segment belonging to a *Yadokarivirus* that encodes a coat protein has not been identified. Alignment of the RNA-seq reads to the genomes of AhoYV1 and PdYV1 reveals an average coverage per nucleotide of 151x and 23x respectively (Fig 9B-C).

A phylogenetic analysis of the RdRP protein sequence reveals a well supported clade that is most closely related to the *Caliciviridae* and separate from other known mycoviral families including *Hypoviridae* and fusariviruses (Fig 9A). AhoYV1 in particular is most closely related to AfSV2; both viruses were isolated from a member of the genus *Aspergillus* [74].

As percent identity between virus RdRP sequences is a common criteria for defining new groups of viruses, the RdRP motif sequences from the viruses in Fig 9A were analyzed to identify a potential threshold value. Indeed a threshold between 26 to 35% identity easily classifies these viruses together, but separate from other known (+)ssRNA viruses (Fig 9D). An

additional threshold of 95% would distinguish each sequence evaluated here as a unique viral species; however, as the pairwise percent identity between AhoYV1 and AfSV2 of 94% is somewhat higher than is generally observed for unique species, a lower threshold may be more appropriate. Identification of additional sequences belonging to this group of viruses will further refine these sequence-based characteristics.

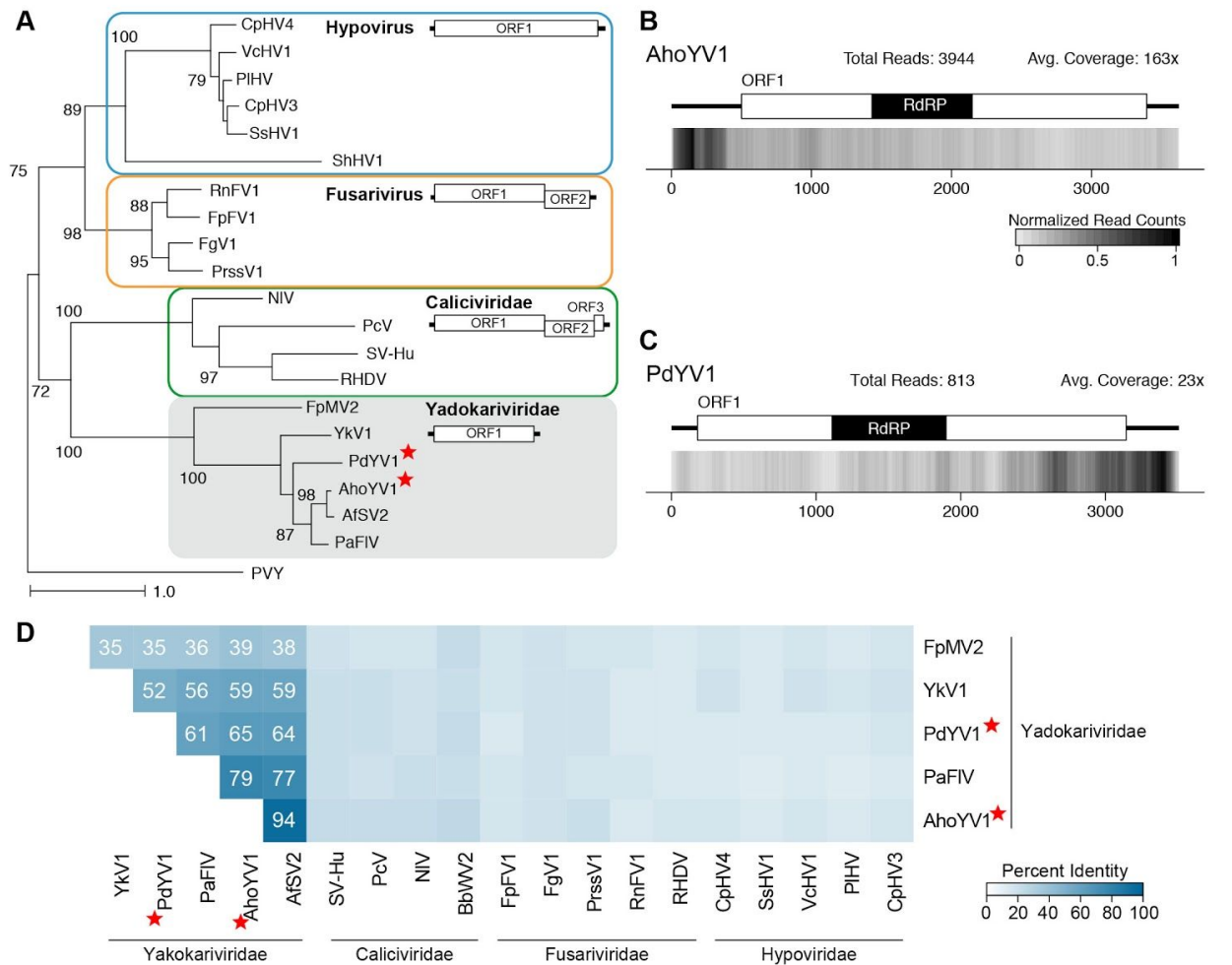


Fig 9. Analyses of *Aspergillus homomorphus yadokarivirus 1* (AhoYV1) and *Penicillium digitatum yadokarivirus 1* (PdYV1) viral genome sequences, members of the proposed family *Yadokariviridae*. (A) A phylogram of the RdRP coding sequence from select (+)ssRNA viruses from the *Yadokariviridae* family, in addition to members of the families *Caliciviridae*, *Fusariviridae*, and *Hypoviridae*; sequence from *Potato Virus Y* (PVY) serves as the outgroup. Virus names have been abbreviated; full names and accession numbers for protein sequences used in the alignment are in Table S2. Phylogenetic tree was generated with RAxML; scale bar represents 1.0 amino acid substitutions per site and numbers at the nodes indicate bootstrap support over 70% (1000 replicates). To the right of the phylogram are representative diagrams of the genome structures for each family of viruses. (B and C) Density plot of read counts per nucleotide across the viral genome for (B) AhoYV1 and (C) PdYV1. Graphical depiction of the genome structure, the predicted single ORF and the RdRP motif are along the top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Legend for (C) is the same as the legend found in (B). Total reads mapped to the virus genome and average read coverage are noted for each virus. (D) Percent identity matrix generated by Clustal-Omega 2.1 using the RdRP motif sequences from the viruses analyzed in (A). For the sake of clarity, only the top half of the matrix is included and the 100% values along the diagonal have been removed. Percent identities greater than 25% are indicated. Newly identified viruses, AhoYV1 and PdYV1, are noted with a red star. RdRP: RNA-dependent RNA Polymerase; ORF: Open Reading Frame.

Another feature common to these viruses is the presence of at least one additional virus within the fungal host. Co-incident infection is not an uncommon occurrence in fungi; however, it has been demonstrated for two yadokariviruses, AfSV2 and YkV1, that isolated virus particles contain a second virus, related to Totiviruses [74,75]. Further, Zhang and colleagues demonstrated that YkV1 is *trans*-encapsidated with the capsid protein from the other virus, YnV1 [75]. As caliciviruses and totiviruses share a distant but strongly supported ancestry following diversification of the picorna-like superfamily [76], and both AfSV2 and YkV1 are present within virions encapsidated by totivirus capsid proteins, it is intriguing to speculate about the mechanisms at play, and the requirements for such an interaction. A totivirus was also identified in the *A. homomorphus* sample (AhoTV1, discussed above), however only a partitivirus (PdPV1-HSF6) was identified along with PdYV1. For the remaining two members of this group of viruses, it is not possible to know the identity of the putative coat-protein donor, as PaFIV1 was found to co-infect *P. aurantiogriseum* along with a fusarivirus, totivirus, unclassified virus, and partitivirus [4] and FpMyV2 was identified in a fungal strain that contained 13 other viruses from a variety of families [38].

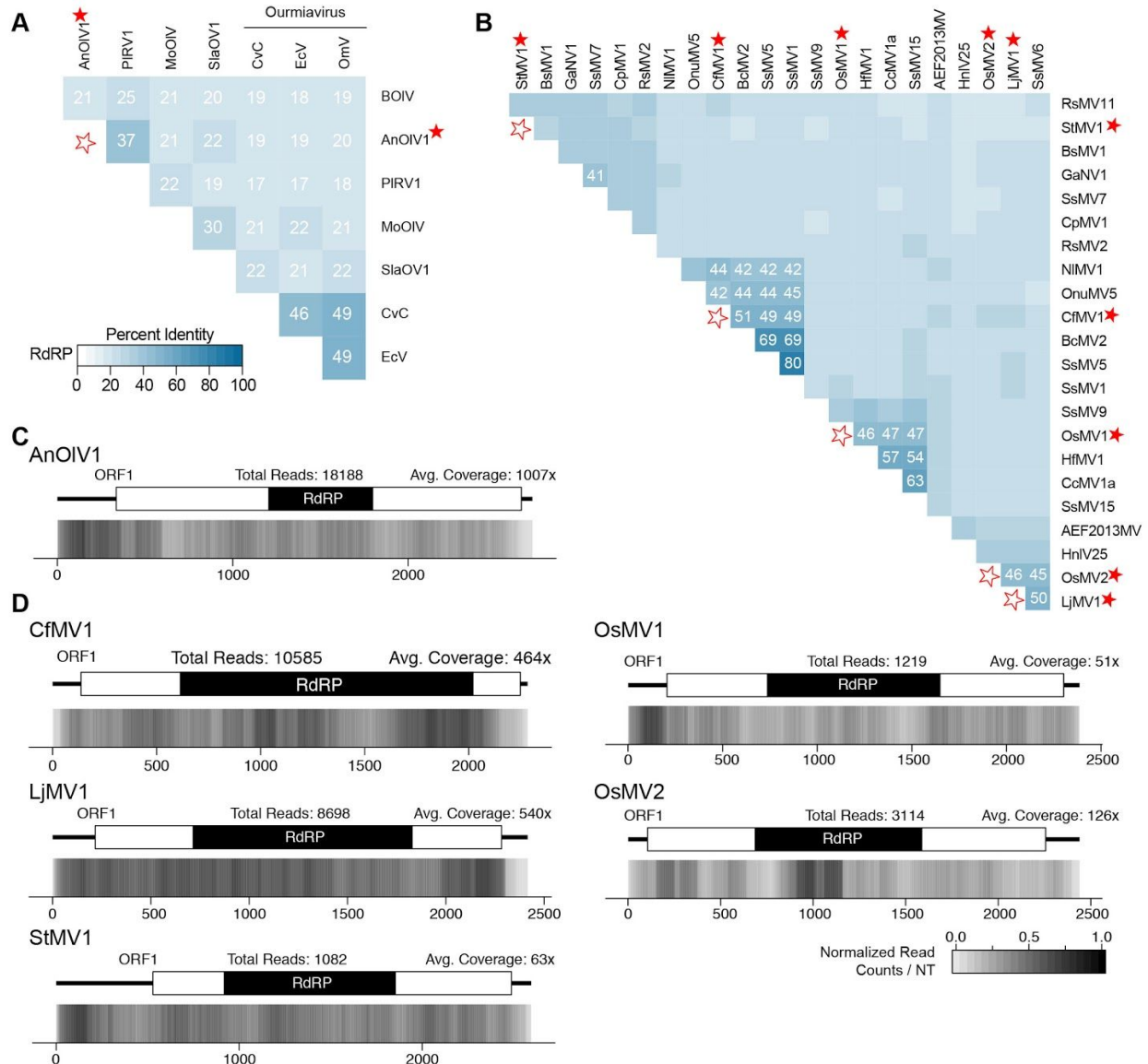
Further study of the fungi that host a member of this group of viruses will answer questions regarding the relationship between the capsid-less virus and the capsid-donor, and may provide an interesting system for studying the evolution of a capsid-less virus. Due to the distinct phylogenetic relationship between these six identified and known (+)ssRNA viruses, along with the unique lifestyle, we support the recommendations to create a new virus family, named *Yadokariviridae* after YkV1, where *yadokari*, is the Japanese word for hermit crab [73,77].

Narnaviridae and Ourmia-like virus

Members of the family *Narnaviridae* have single molecule, positive-strand RNA genomes ranging in length from 2.3 to 2.9 kb and encode an RdRP as the only polypeptide [78]. There are currently two genera within this family: *Narnavirus* and *Mitovirus*, where mitoviruses are characterized as viruses that replicate within the mitochondria of filamentous fungal hosts [78]. Indeed, a non-interrupted ORF is only predicted in these viruses when using the mitochondrial codon usage where tryptophan is encoded by UGA [79]. Here we describe five new viruses with similarity to members of the genus *Mitovirus* from RNA-seq datasets from four fungal hosts: *Setosphaeria turcica mitovirus 1* (StMV1) (PRJNA250530; unpublished), *Colletotrichum falcatum mitovirus 1* (CfMV1) (PRJNA272832; [80]), *Loramyces juncicola mitovirus 1* (LjMV1)

(PRJNA372853; unpublished), and *Ophiocordyceps sinensis mitovirus 1 and 2* (OsMV1, OsMV2) (PRJNA292632; [81]).

The five mitoviruses have genome sizes within the expected range (Table 1), and a single ORF is present only when using the mitochondrial codon table. The percent of A+U in the coding strand for these five viruses ranges from 53% to 70%; characteristically, mitoviruses have 62 to 73% A+U [78], although the percent observed in recently identified viruses is as low as 53% [5]. The predicted protein within each genome contains a Mitovir_RNA_pol domain, with



S8 Fig. Analyses of identified viruses related to *Ourmiavirus* and *Narnaviridae*. (A and B) Percent identity matrix of the RdRP protein coding sequence, generated by Clustal-Omega 2.1, of ourmia-like mycoviruses (A) and the genus *Mitovirus* from the family *Narnaviridae* (B). For the sake of clarity, only the top half of the matrix is shown and the 100% identity along the diagonal has been removed; outlined stars are used to note junction point for the row and column of the virus(es) identified. The heatmap scale in (B) is the same as found in (A). Sequences used in analysis can be found in Table S2. (C and D) Density plot of read counts per nucleotide across the viral genomes of the ourmia-like viruses (C) and mitoviruses (D) identified in this study. Graphical depiction of the genome structure, ORF, and predicted RdRP motif for each genome are along the top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. Viruses identified in this study are noted with a red star. RdRP: RNA-dependent RNA Polymerase.

a length ranging from 652 to 716 aa. Alignment of the RNA-seq reads to the viral genomes demonstrates end to end coverage for each virus, with average coverage per nucleotide ranging from 51x for OsMV1 to 540x for LjMV1 (S8D Fig).

A phylogenetic analysis of the RdRP coding sequence demonstrates strong support for all five newly identified viruses among members of the genus *Mitovirus* (Fig 7B). Criteria for species demarcation within the genus *Mitovirus* is not yet full defined, however, viruses with greater than 90% identity within the RdRP protein sequences are considered the same species [20]. Here we observe that StMV1, CfMV1, OsMV1, OsMV2, and LjMV1 have less than 52% identity to any currently available mitovirus sequence (S8B Fig). Taken together, the above characteristics demonstrate that these five viruses are new species within the genus *Mitovirus*.

Ourmiaviruses are also positive-strand RNA viruses, composed of three genome segments that encode an RdRP, movement protein, and coat protein, isolated from plant hosts [19,82]. Despite usage of standard eukaryotic codons, phylogenetically the RdRP protein sequence is most closely related to that of *Narnavirus* [19]. Recently viruses have been identified from fungal hosts with sequence similarity to that of the plant-hosted ourmiaviruses. One such virus was identified in this study: *Aspergillus neoniger ourmia-like virus 1* (AnOIV1) (PRJNA250996; unpublished). The single segment, 2.7 kb genome encodes a single polyprotein that contains an RdRP motif (Table 1). Alignment of the RNA-seq data to the genome revealed an average coverage of 1007x (S8C Fig).

A phylogenetic analysis of this virus, along with related ourmia-like viruses from other fungi, plant-hosted ourmiaviruses, and mitoviruses illustrates a well supported clade containing AnOIV1 along with *Phomopsis longicolla RNA virus 1* (PIRV1) [83] and *Botrytis ourmia-like virus HAZ2-3* (BoIV) [84] (Fig 7B). A percent identity analysis of the fungal and plant viruses found in Fig 7B demonstrates a percent identity of approximately 50% among the recognized *Ourmiaviruses*, while the fungal-hosted viruses shared from 17 to 37% identity (S8A Fig), with AnOIV1 and PIRV1 the most similar. As there is limited sequence similarity, and AnOIV1 was identified from a new fungal host, it would appear that this is a new species within this yet-to-be fully characterized viral genus. Identification of additional *Ourmiaviruses* and related mycoviruses will shed further light on the specific relationship between these plant and fungal viruses.

Hypoviridae

The *Hypoviridae* family consists of positive sense, single-stranded RNA viruses that are approximately 9.1 to 12.7 kb in length, containing either one long ORF or two ORFs [85]. Further, these viruses are capsidless and use a host-derived, trans-Golgi network for genome replication [86,87]. Infection by Hypoviruses can often result in reduced fungal virulence, a hallmark symptom of early members of the *Hypoviridae* family.

Hypoviruses identified from four transcriptomic datasets are described here: *Sclerotinia homoeocarpa hypovirus 1* (ShHV1) (; PRJNA167556[88]), *Trichoderma asperellum hypovirus 1* (TaHV1) (PRJNA261111; [89]), *Setosphaeria turcica hypovirus 1* (StHV1) (PRJNA250530; unpublished), and *Fusarium graminearum hypovirus 1* (FgHV1) (PRJNA263651; [90]). The virus FgHV1 was expected, as the sequencing experiment featured *F. graminearum* strain HN10 infected with FgHV1 [90]; to our knowledge, this is the first mycovirus identified from *S. turcica* and *T. asperellum*, and the first hypovirus from *S. homoeocarpa*.

There are currently four species officially recognized in the *Hypoviridae* family: *Cryphonectria hypovirus 1, 2, 3, and 4* (CHV1, CHV2, CHV3, and CHV4), although several other related viruses have recently been identified from eight other fungal species. Of the currently recognized hypoviruses two clades form that appear to be distantly related, whereby CHV1 and CHV2 are most closely related and CHV3 and CHV4 are together in a separate grouping. To this end, two new genera have been proposed, *Alphahypovirus* [91] and *Betahypovirus* [92], where CHV1 and CHV2 are classified in the former and CHV3 and CHV4 in the later [85]. A phylogenetic analysis using the RdRP protein sequence of the four viruses identified here, along with known and related hypoviruses, revealed that FgHV1, ShHV1, and TaHV1 groups with the *Alphahypoviruses* while StHV1 was found with the *Betahypoviruses* (Fig 7C). A further analysis of these viral sequences, described below within the context of the two recently proposed genera, determined that ShHV1, StHV1, and TaHV1 are new species within the *Hypovirus* genus.

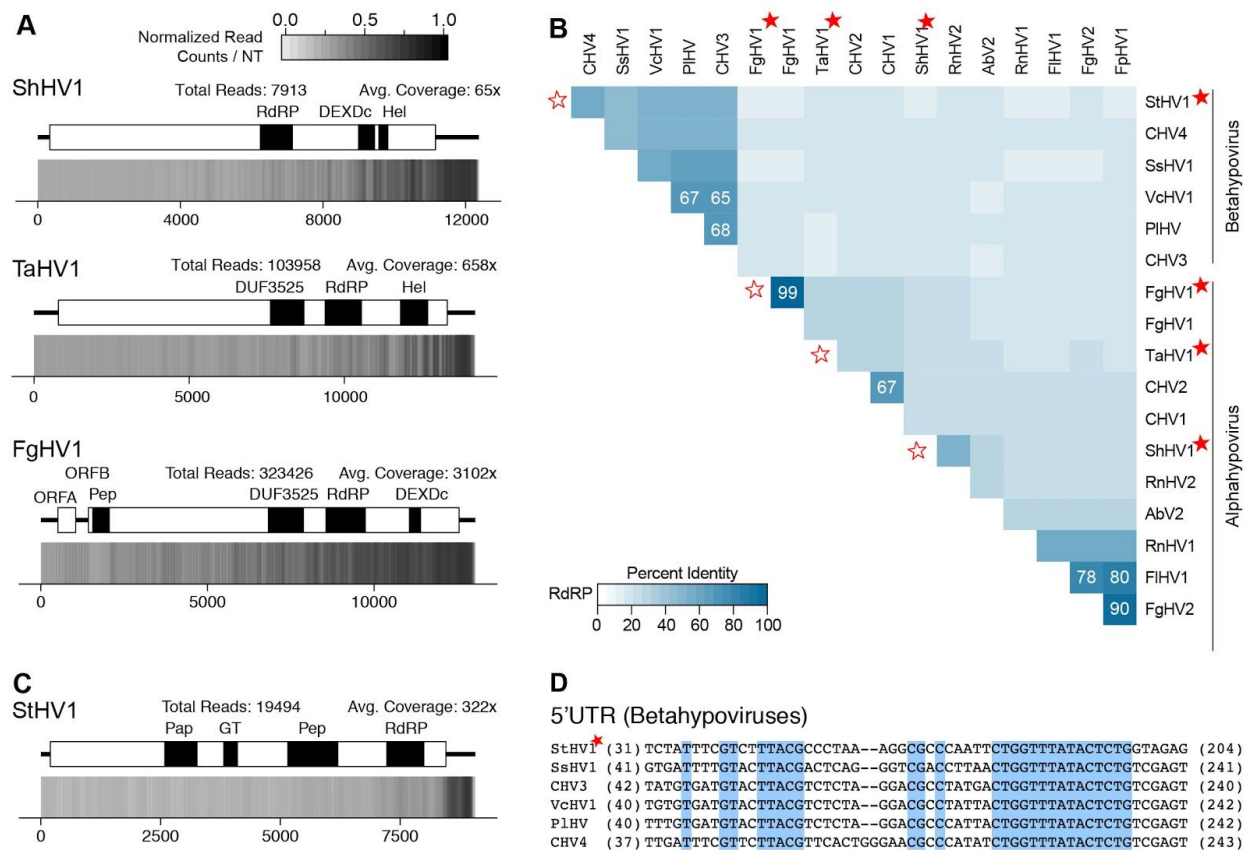
Alphahypovirus

Initially, members of *Alphahypovirus* were characterized as having two ORFs, where the second ORF contained protease, RdRP, and Helicase domains. This included CHV1 [93], CHV2 [94], FgHV1 [91], and *Rosellinia necatrix hypovirus 2* (RnHV2) [95]. However, identification of additional viruses has revealed members with a single, large ORF, including *Fusarium langsethiae hypovirus 1* (FIHV1) [96], *Fusarium graminearum hypovirus 2* (FgHV2) [91],

Fusarium poae hypovirus 1 (FpHV1) [38], and *Rosellinia necatrix hypovirus 1* (RnHV1) [95].

Independent of the number of ORFs, alphahypoviruses have large genomes, with lengths ranging from 12.5 to 14.4 kb. Similarly, newly identified viruses ShHV1 and TaHV1 have genome lengths of 12.4 and 14.2 kb respectively, and single-ORF genome structures (S9A Fig, Table 1).

For a new species to be classified, criteria generally include host species and percent identity with other identified species. Official criteria do not yet exist for the family *Hypoviridae*, and separate criteria for the two proposed genera may be appropriate given the lack of sequence homology between the two groups. Previously it has been observed that CHV1 and CHV2, while both isolated from *Cryphonectria parasitica*, demonstrate only ~67% percent



S9 Fig. Analyses of identified viral genomes belonging to the family *Hypoviridae*. (A and C) Density plot of read counts per nucleotide across the viral genomes that group with betahypoviruses (A) and alphahypoviruses (C). Graphical depiction of the genome structure, predicted ORFs, and encoded motifs are along the top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. (B) Percent identity matrix, generated by Clustal-Omega 2.1, of the RdRP-motif containing protein coding sequence from the newly identified viruses and related hypoviruses. For sake of clarity, the 100% identity along the diagonal has been removed and outlined stars are used to note the junction point for the row and column of the viruses identified. Where sequences have similarity $\geq 60\%$ is indicated. Full names and accession numbers for sequences used are in Table S2. (D) Alignment of 5'UTR sequences from alphahypoviruses where conserved nucleotides are highlighted in blue and amino acids not included in the alignment are noted in parentheses. Viruses identified in this study are noted with a red star. Pap: papain-like cysteine protease domain homologs; RdRP: RNA-dependent RNA Polymerase; Hel: Helicase; GT: UDP-glucose/sterol glucosyltransferase; Pep: peptidase; DUF: Domain of Unknown Function.

identity within the RdRP amino acid sequence. Expanding this analysis to all viruses within the genus *Alphahypovirus* reveals a maximal identity of 90%, between FgHV2 and FpHV1 (S9B Fig); this excludes the 99% identity observed between the two sequences of FgHV1 which are known to be the same species of virus [90]. Interestingly, there is a cluster of similar hypoviruses from three different *Fusarium* species, FIHV1, FpHV1, and FgHV2, which excludes the fourth *Fusarium* hypovirus, FgHV1. An appropriate threshold for evaluating new species of *Alphahypovirus* remains to be determined; however as ShHV1 and TaHV1 share only 50% and 31% identity with RnHV2 and CHV1 respectively, and were identified from unique fungal hosts, these viruses appear to be new species within the family *Hypovirus*.

A final analysis of the three alphahypoviruses identified from this study was the alignment of RNA-seq reads to the viral genomes. Total reads that aligned range from 7,913 for ShHV1 to 323,433 for FgHV1, and in all three instances, the peak coverage is closest to the 3' end of the genome (S9A Fig).

Betahypovirus

Viruses in the *Betahypovirus* genus generally encode one ORF that contains protease, RdRp, Helicase, and UGT domains [92] and tend to have shorter genomes than the *Alphahypoviruses*, at approximately 9 to 11 kb in length. These same characteristics are also observed in StHV1, which has a length of 9163 bp and a polyprotein predicted to encode a papain-like cysteine protease domain along with a UGT, peptidase-C97, RdRP, and Helicase-HA2 domain (S9C Fig, Table 1). Additionally, the 5' end of the 5'UTR sequence has high sequence identity across viral genomes from this genus; a multiple sequence alignment of this region from StHV1 and related *Betahypoviruses* reveals similarly high conservation within the StHV1 genome, including a 15-nt sequence that is present in all viruses analyzed (CTGGTTTATACTCTG) (S9D Fig).

Characterization of the StHV1 sequence also included alignment of the RNA-seq reads to the viral genome, resulting in a total of 376,89 reads aligned, and an average coverage per nucleotide of 617x (S9C Fig).

A percent identity analysis of the RdRP amino acid sequences of viruses within the genus *Betahypovirus* reveals a maximal identity of 68%, between PIHV and CHV3 (S9B Fig) while StHV1 shares 46% to 54% identity with other viruses. As such, a threshold of 70% identity within the RdRP amino acid is suggested as a criteria to distinguish between different species of *Betahypoviruses*.

Fusariviridae

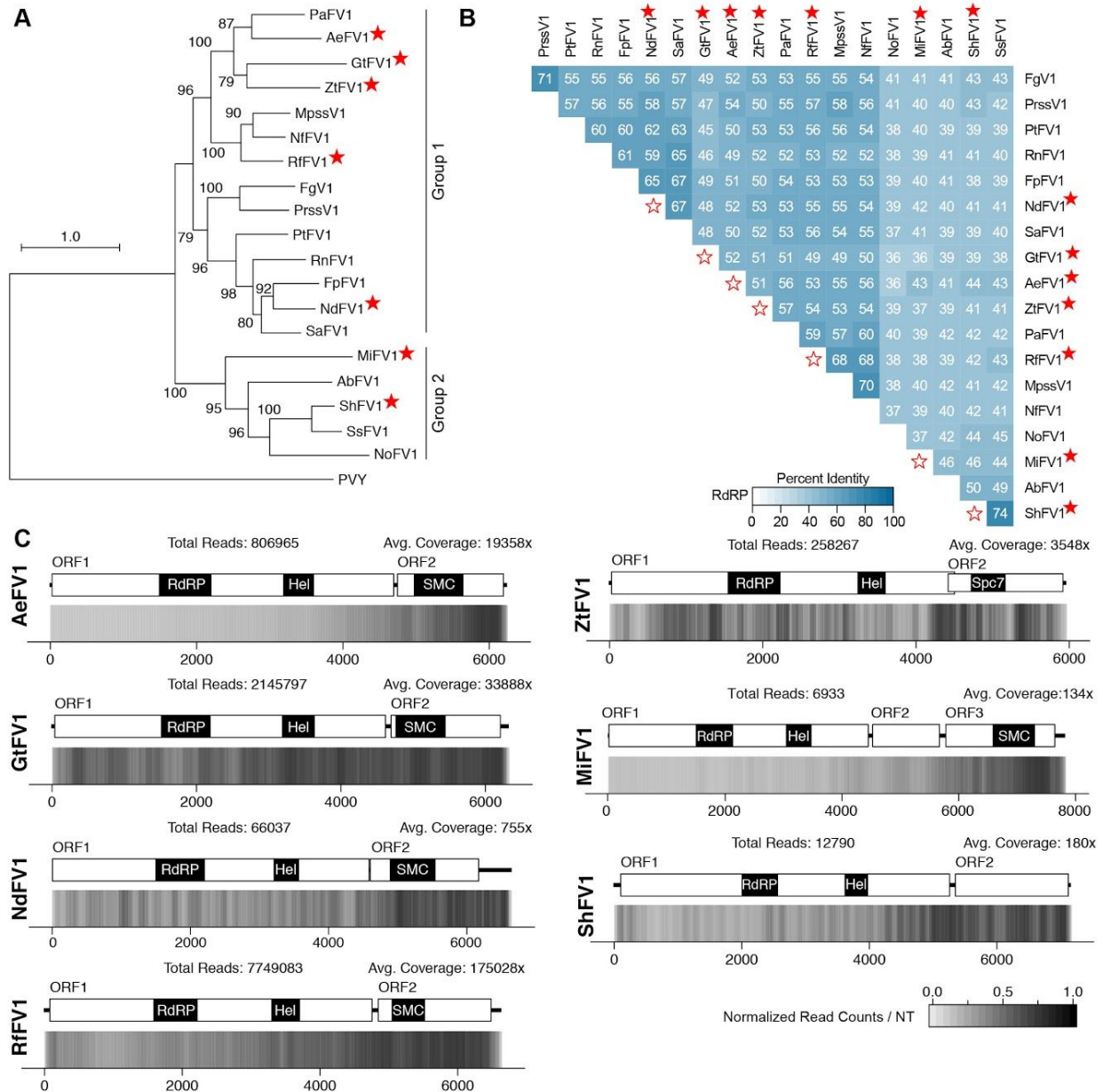
Members of the proposed family *Fusariviridae* have mono-segmented, plus strand, single-stranded RNA genomes that encode two to four ORFs. ORF1 is the largest and contains the RdRP and Helicase motifs, while the subsequent and second largest ORF encodes an unknown protein, often containing an SMC motif. One to two additional small proteins have been predicted at the 3' end of some virus genomes. Here we have identified seven new viruses that group with fusariviruses.

Aspergillus ellipticus fusarivirus 1 (AeFV1) (PRJNA250911; unpublished), *Morchella importuna fusarivirus 1* (MiFV1) (PRJNA372858; unpublished), *Neurospora discreta fusarivirus 1* (NdFV1) (PRJNA257829; [97]), and *Rutstroemia firma fusarivirus 1* (RfFV1) (PRJNA372878; unpublished) are to our knowledge, the first viruses to be identified in these species. *Sclerotinia homoeocarpa fusarivirus 1* (ShFV1) (PRJNA167556; [88]) and *Gaeumannomyces tritici fusarivirus 1* (GtFV1) (PRJNA268052; [34]) are the first fusariviruses identified in these fungi, and this is the first full-length description of *Zymoseptoria tritici fusarivirus 1* (ZtFV1) (PRJNA237967; [98]) (PRJEB8798; [99]) (PRJNA253135; [100]) (PRJNA179083; unpublished).

The genome of MiFV1 encodes three predicted ORFs, where the first contains both the RdRP and helicase motifs, the second contains no predicted domains, and the third ORF contains an SMC (structural maintenance of chromosomes) domain (S10C Fig). The genomic structure of the other six viruses identified in this study encode two predicted ORFs, again with the first encoding RdRP and helicase motifs (S10C Fig). ORF2 of ZtFV1 contains a predicted Spc7 domain (found in kinetochore proteins), a domain also found in *Rosellinia necatrix fusarivirus 1* (RnFV1) [101]. Within ORF2 of AeFV1, GtFV1, NdFV1 and RfFV1 is a predicted SMC domain, while no functional domain is predicted within ORF2 of ShFV1. A previous phylogenetic analysis determined that the closest relatives to the viral SMCs are the proteins from the eukaryotic SMC5 and SMC6 subfamily and that because the SMC domain is widely present throughout members of the *Fusariviridae* family, a common ancestral origin is probable [102].

A phylogenetic analysis of the RdRP amino acid sequence of these seven viruses and twelve other published fusariviruses reveals two strongly supported clusters (S10A Fig). "Group 1" contains 14 viruses while "Group 2" contains the remaining five fusariviruses. The genome structures of the two clusters suggests some general patterns; all members of Group 1 have a predicted functional domain within the second largest ORF, either an SMC or an Spc7 domain as described above. Although, FgV1 encodes three ORFs, unlike related viruses that encode

just two. Conversely, viruses within Group 2 generally do not have a functional domain predicted within ORF2, the exception is MiFV1 which contains an SMC domain. Further, three ORFs are predicted for four of the five members, with ShFV1 being the exception. Identification



S10 Fig. Analyses of viral genomes that belong with known fusariviruses. (A) Phylogram of fusariviruses based on the ORF1 protein coding sequence which contains the RdRP motif. Full names and accession numbers for protein sequences used are in Table S2, including the outgroup Potato Virus Y (PVY). Scale bar represent 1.0 amino acid substitutions per site and numbers at the nodes indicate bootstrap support over 70% (1000 replicates). (B) Percent identity matrix generated from Clustal 2.0 using the RdRP motif sequences from fusariviruses. For sake of clarity, the 100% identity along the diagonal has been removed and open stars note the junction point between the row and column of the viruses identified in this study. (C) Density plot of read counts per nucleotide across the viral genomes identified in this study. Graphical depiction of the genome structure, predicted ORFs and encoded motifs are on top and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count for a given genome equals 1. Total reads mapped to the virus genome and average read coverage are noted. RdRP: RNA-dependent RNA Polymerase; Hel: Helicase domain; SMC: SMC N terminal domain; Spc7: Spc7 kinetochore domain.

of additional viral genomes that belong to this family will help to elucidate the support for these separate clusters and may reveal additional distinguishing characteristics.

As *Fusariviridae* is not yet an officially recognized family of viruses, specific characteristics for distinguishing new viral species or genera have yet to be established. In addition to phylogenetic analyses and identity of fungal host species, percent identity across the protein coding sequence of the RdRP-containing ORF is frequently used as a classifier. To date, fusariviruses previously characterized have all been isolated from unique fungal species, including two different species of *Fusarium*, and none share greater than 57% identity with one another (S10B Fig). Viruses identified in this study are also from unique fungal species, including ShFV1 which was isolated from a different species than *Sclerotinia sclerotiorum fusarivirus 1* (SsFV1) [103]. Using the data available to date, two criteria are recommended to qualify viruses as new species: a unique fungal host and a percent identity of the RdRP sequence $\leq 60\%$ in comparison to existing fungal sequences.

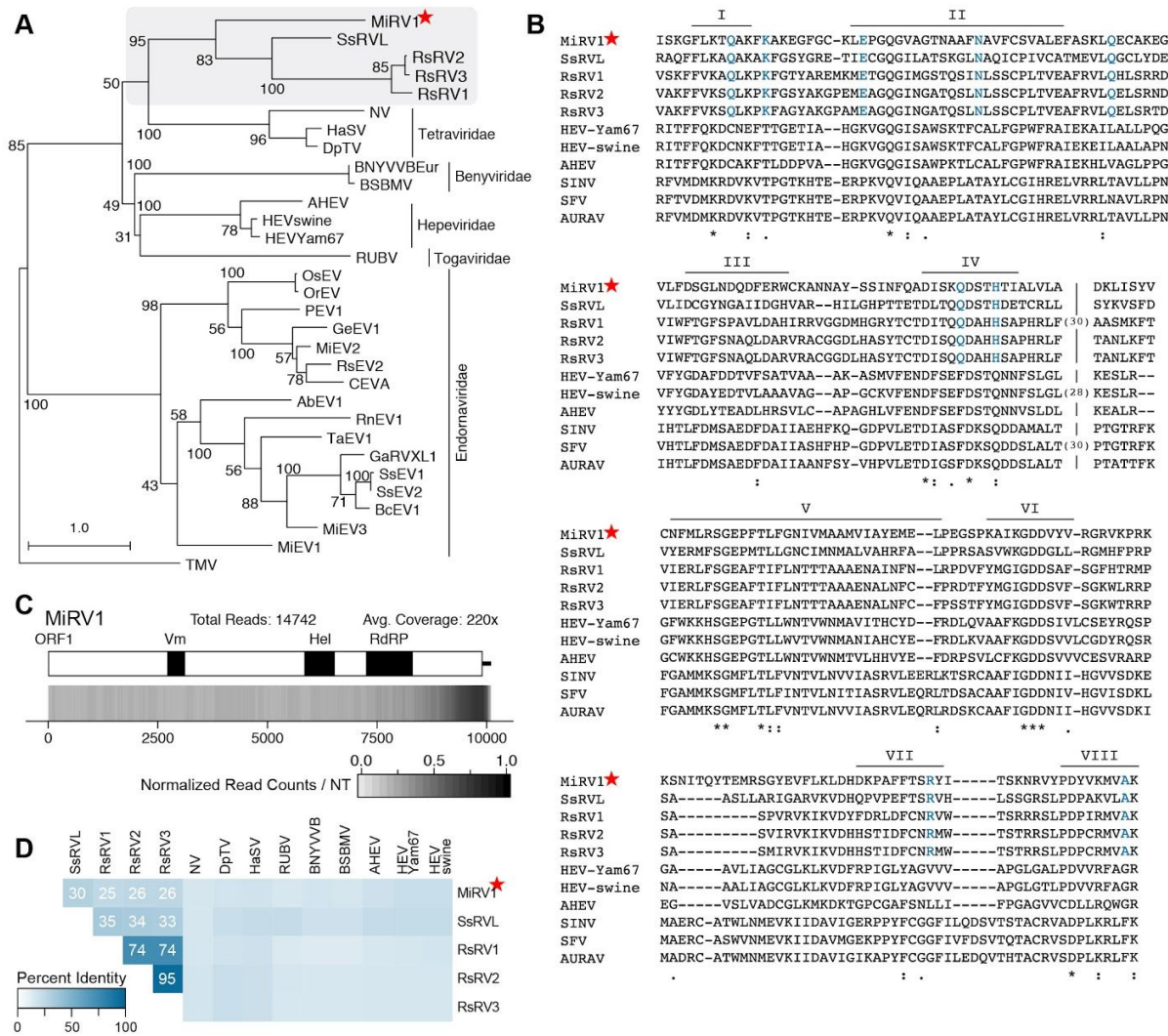
The ZtFV1 sequence described here was identified within the SRA sample PRJNA179083, however interestingly there are three other RNA-seq experiments available for this organism from three other research groups. All four sequencing samples appear to contain the same fusarivirus; indeed a full length sequence was isolated from samples from both the Max Planck Institute for Terrestrial Microbiology (PRJNA237967; [98]) and Rothamsted Research (PRJEB8798; [99]). For the sample from the Institut national de la recherche agronomique (INRA) (PRJNA253135; [100]), a contig covering 76% of the full length was identified, and manual analysis determined sufficient reads to support the full length sequence. The only difference between the four sequences was the length of 5' poly-G and 3' poly-A sequences, as such, the most parsimonious version was selected and deposited into GenBank. A bowtie2 alignment of the reads from each sample to this final viral genome sequence did not identify any SNPs for any sample, although some regions of heterozygosity are present among the samples. Nonetheless, the average read coverage per nucleotide using reads from each sample individually was high and ranged from 3548x to 15278x (S10C Fig). Similarly, average coverage per nucleotide was robust for the other fusariviruses identified, ranging from 180x for ShFV1 to 33888x for GtFV1 (S10C Fig).

MiRV1

In addition to the three endornaviruses and one fusarivirus described above, a fifth viral sequence was identified in the *M. importuna* sample (PRJNA372858; unpublished). The

single-segment sequence was determined to be 10,132 nt in length, with a single predicted ORF of 3,295 aa (Table 1). A conserved domain search revealed a viral methyltransferase domain and a viral helicase domain in addition to the RdRP_2 motif similar to positive-strand RNA viruses (S11C Fig). A BLAST-P search with the predicted protein sequence returned a single full length mycovirus sequence, *Sclerotinia sclerotiorum RNA virus L* (SsRV-L) [104], in addition to matches to Hepatitis Virus E and Cordoba virus. Other top hits include partial sequences corresponding to the RdRP domain of *Rhizoctonia solani RNA virus 1, 2, and 3* (RsRV-1, RsRV-2, RsRV-3) [105]. Both MiRV1 and SsRV-L contain a single ORF with methyltransferase, helicase and RdRP domains, and neither contain peptidase motifs [104].

Previous phylogenetic analyses of SsRV-L and the three viruses from *R. solani* demonstrated a relationship between these viruses and members of the alpha-like virus superfamily, specifically viruses from the *Hepeviridae* and *Alphatetraviridae* families [104,105]. *Endornaviridae*, whose members infect both fungi and plants, are also a family of alpha-like viruses [106], but appear to be more distantly related to this emerging group of mycoviruses. Expanding the phylogenetic analysis to include MiRV1 confirms that MiRV1 is a member of this new clade of mycoviruses (S11A Fig). Further, while a multiple sequence alignment of the RdRP motif sequences demonstrated the presence of the eight conserved RdRP motifs common among (+)ssRNA viruses, certain residues within the motifs were shared only among the mycoviruses that group with MiRV1 (S11B Fig). The low percent identity between MiRV1 and the related mycoviruses indicates that this is a new species of virus (S11D Fig). Identification of additional related viral sequences will help elucidate the relationship of MiRV1 and the rest of the Alpha-like viruses.



S11 Fig. Sequence-based analyses of *Morchella importuna* RNA virus 1 (MiRV1). (A) Phylogram of the RdRP-motif for MiRV1, related mycoviruses, and Alpha-like viruses. Full names and accession numbers for protein sequences used are in Table S2, including the outgroup Tobacco Mosaic Virus (TMV). Scale bar represent 1.0 amino acid substitutions per site and numbers at the nodes indicate bootstrap support over 70% (1000 replicates). (B) Alignment of the RdRP motif from MiRV1 and related sequences. Conserved motifs within (+)ssRNA RdRP sequences are noted (I to VIII) and amino acid residues found only in the five related mycoviral sequences are noted in blue. Numbers in brackets are amino acids excluded from the alignment. (C) Density plot of read counts per nucleotide across the MiRV1 genome. Graphical depiction of the genome structure, predicted ORF, and encoded motifs are along the top, and a heatmap of normalized read counts is on the bottom. Read counts are normalized to a 0 - 1 scale, where the maximum read count equals 1. Total reads mapped to the virus genome and the average read coverage are noted. (D) Percent identity matrix generated from Clustal 2.0 using the RdRP motif sequences from MiRV1, mycoviral relatives and select Alpha-like viruses. For sake of clarity, the 100% identity along the diagonal has been removed and only identities $\geq 25\%$ are indicated. Red star notes MiRV1, the virus identified in this study. RdRP: RNA-dependent RNA Polymerase; Vm: Viral methyltransferase; Hel: Helicase.

Unknown comovirus-like sequences

An RdRP-containing contig with a top match to Turnip Ringspot Virus (TRV) was identified in sequencing samples from *Colletotrichum tofieldiae* (;PRJNA287627 [107]) and *Zymoseptoria brevis* (PRJNA277175; [108]). Further inquiry also identified the second genomic segment of

TRV, which encodes the movement and coat proteins, from both the *C. tofieldiae* and *Z. brevis* Trinity contigs. The coding sequences of RNA1 and RNA2 are 98.5% and 99.7% identical, respectively, suggesting a close relationship between the sequences, despite the unrelated hosts.

As Turnip Ringspot Virus (genus: *Comovirus*; family: *Secoviridae*) can infect *Brassicaceae* [109], and a fungal host has not yet been described for viruses from this family, we postulate that this sequence may be due to cross contamination. The BioProject containing the *C. tofieldiae* sample included both media-grown fungi and an infection series of *Arabidopsis thaliana*. Further analysis of the infection series determined that the viral sequence is present in 19 of 24 (79%) of the mock-infected *A. thaliana* samples, 19 of 24 (79%) of the *A. thaliana* plus *C. tofieldiae* samples, and two of three fungus-only samples. An explanation for the uneven distribution of the viral sequences among these samples remains to be determined. Further, the *Z. brevis* sample was part of a media-grown, fungus-only transcriptomic experiment and as such, a putative source of the TRV sequences is not readily apparent. The *Z. brevis* sample is from Christian-Albrechts University of Kiel (Kiel, Germany) and was sequenced at the Max Planck Genome Center (Cologne, Germany) while the *C. tofieldiae* sample is from the Max Planck Institute for Plant Breeding Research (Kohn, Germany) and was also sequenced at the Max Planck Genome Center Cologne (Cologne, Germany). At this junction, it is not possible to determine the host or source of the viral genome sequences identified in the *C. tofieldiae* and *Z. brevis* samples.

Summary and Conclusions

Here we describe the power of a systematic approach for analyzing public RNA-seq datasets to identify RNA virus genomes. Application of this bioinformatics pipeline to 569 transcriptomic datasets from 312 species of Pezizomycotina fungi identified 59 full-length viral genomes, 53 of which were determined to be new viruses.

The success of the pipeline was due to (a) the ability to fully assemble a viral RdRP coding sequence from as few as 136 reads (or 12x coverage) and (b) the strong signature of the viral RdRP domain. Benefits of this analysis include suitability to either rRNA-depleted total RNA or poly-A purified RNA. As viral particles enrichment is not required, which would select against non-particle forming viruses, virus identification via RNA-seq data can easily be determined from regular fungal transcriptomic experiments. Notable limitations of this approach include bias for

viruses with a recognizable RdRP domain, which automatically excludes DNA viruses and may select against (-)ssRNA viruses. Further, robust analyses are required following the identification of putative viral sequences to confirm homology and relatedness to the suite of known viruses.

This analysis is not just an academic exercise in computational pipeline development and execution. It also works to answer the biological question, what fungi are infected by what viruses? It is becoming increasingly clear that there is a dynamic range of responses due to viral infection, outside of the well documented asymptomatic or hypovirulence phenotypes, that reflect the complex relationships present during multipartite interactions. A recent review by Roossinck [110] highlights a number of mutualistic or symbiogenic relationships between viruses and their hosts. This includes killer yeasts: yeast that are infected with a virus which produces toxins that kill competitors to the host fungi while conferring resistance to the host [111]. Another example is the tripartite interaction between the tropical panic grass *Dichanthelium lanuginosum*, the endophytic fungi *Curvularia protuberata* and *Curvularia thermal tolerance virus*, the virus that infects *C. protuberata*. Thermal tolerance to high soil temperatures is only conferred to the plant when the fungi is present and is infected with the virus [112]. Another phenotype, different from canonical hypovirulence, is the reduction in fungicide resistance observed following viral infection of prochloraz-resistant strain of *Penicillium digitatum* [113].

Multipartite relationship also include virus-virus interactions within a host as many fungi are concurrently infected by more than one mycovirus; here we identified over 10 such fungi. Recently, it was described that a positive effect on accumulation of a viral genome was observed when a second, unrelated virus was present in *Rosellinia necatrix* [75]. Conversely, coinfection of *C. parasitica* by a totivirus and a suppressor-deficient hypovirus lead to induction of antiviral defense genes and ultimately inhibition of totivirus replication [114].

Indeed asymptomatic or latent infections can downplay the active antiviral processes at work within the fungus and the important role a successful antiviral defense has on fungal phenotypes. Small RNA-mediated antiviral defense systems, crucial for preventing a deleterious effect on fungal health, have been identified in *Cryphonectria parasitica* [115], *Colletotrichum higginsianum* [8], and *Sclerotinia sclerotiorum* [116]. Further, it can be assumed that antiviral defenses have contributed to the under-identification of mycoviral infections in fungi, including many of the fungi identified here that are infected with a virus, but lack studies into RNA silencing.

This systematic analysis also highlights the uneven distribution of sequencing data among members of the Pezizomycotina subphylum. While strides are being made, particularly due to large-scale approaches such as the 1000 Fungal Genomes Project, a fuller understanding of the mycoviral landscape and therefore the effect a virus has on fungal phenotype requires continuous inquiry. As the number of Pezizomycotina fungi is estimated to be greater than 32,000 [117] and the current scope of known fungal hosts is less than 500 species, it stands to reason there remains a wide range of viral diversity yet to be identified and characterized. Further, this study focused solely on the subphylum Pezizomycotina, meaning members of the sub-phyla Saccharomycotina and Taphrinomycotina and the unranked group Ascomycota incertae sedis are not included. Beyond the fungal kingdom, what about datasets from the animal and plant kingdoms? While the computational requirements are large to investigate these tens of thousands of RNA-seq datasets, the rate limiting steps are the manual curation required at the start of the pipeline, to analyze metadata and identify suitable samples, and at the end of the pipeline, to evaluate and characterize the identified viral genomes. Other bioinformatic approaches under consideration are (a) downsampling the reads to a suitable sized input for the pipeline and / or (b) direct translation of the raw reads followed by analysis for RdRP domain signatures. Finally, sequencing technologies that produce reads >1 kb in length provide an even more direct route to virus detection as entire viral genomes could be sequenced as single reads. For all future RNA-seq experiments, we highly recommend that the data be analyzed for the presence of viral RdRP signatures as part of any routine bioinformatic pipeline.

Materials and Methods

Identifying publicly available fungal RNA-seq datasets

Programs available via the Entrez Programming Utilities (“E-utilities”) were used to search the NCBI Short Read Archive (SRA) for datasets that matched the search criteria. For ease of searching, the fungal kingdom was divided into three phylum level searches: ascomycota (subphylum Pezizomycotina), basidiomycota, and other. General search criteria used for all phylum-level searches included the phrase “NOT LS454 NOT PACBIO NOT ABI_SOLID AND biomol rna [PROP]” to identify RNA-seq datasets from Illumina (or Solexa) instruments. The

E-utilities commands `esearch` and `efetch` combined to create a comma-separated table of all matching SRA records.

Pipeline for identification of viral-like sequences

The output file from the E-utilities search command was manually parsed to create a unique list of all BioProjects and fungal species. Multiple fungal species found under one BioProject number were split into separate entries. Most BioProjects included more than one sample; for example different environmental treatments or wild-type versus mutant. In general, the SRA sample selected for analysis was a wild-type-like strain grown on media. When available, the fungal genome sequence was either downloaded from NCBI or Joint Genome Institute (JGI).

The computational steps for identifying viral sequences start with a downloaded raw reads file from SRA. When a fungal genome sequence was available, RNA-seq reads were aligned to the genome using `bowtie2` (version 2.2.9) [118], and the unaligned reads were captured in a separate file using the flag `--un`. A *de novo* assembly of these unaligned reads was created using `Trinity` (v2.1.1) [119] and resultant contigs were analyzed for open reading frames (ORFs) by `TransDecoder` (version 2.1) [120]. Finally, these predicted protein sequences were queried against a custom database of viral RNA dependent RNA polymerases (RdRPs) using `HMMscan` (version 3.1b2; [121]). In general, when the data was paired end, the data present within pair one was sufficient data for identification of viral contigs. Occasionally read depth from the virus genome was low, resulting in fragmented viral contigs. In these instances, the pipeline was rerun using both pair one and pair two data. Due to the nature of the `--un` flag of `bowtie2`, the pairs were treated as individual single-end datasets, instead of paired-end data. Assembly of the virus from *Trichoderma harzianum* (PRJNA216008), was particularly recalcitrant and as such, all six SRA paired-end datasets were used in the pipeline for identification of ThAV1.

Contig extension and identification

Contigs identified containing putative viral RdRP domains were further analyzed by first aligning the nucleotide sequence to the 'nr' database at NCBI using `BLASTX`. A variety of manual annotation steps were required following confirmation of a viral sequence. Some viral genomes, such as those matching the *Partitiviridae* or *Chrysoviridae* families, were predicted to contain additional viral genes encoded on separate contigs; these ORFs encode for coat proteins and/or hypothetical proteins. S2 Table summarizes the viral RdRP sequences identified and, where

applicable, which known BLASTX-identified sequences were used to query the Trinity assembly for additional contigs.

Occasionally contigs were identified where the predicted ORF(s) extended to either the 5'- or 3'-most edge; this indicated that the sequence in hand was truncated. In these instances, the RdRP sequence from the best BLAST match was used to query the Trinity contigs to identify putative viral sequences outside the known RdRP contig. Any contigs identified were then stitched together to create a final sequence with a full-length ORF. Start and stop codons were predicted using ORFinder at NCBI [122].

Bioinformatic Tools and Analyses

Sankey Diagram

The online tool SankeyMATIC [123] was used generate the Sankey Diagram in Fig 1. Data tables used as input are available through the FigShare FileSet (DOI: 10.6084/m9.figshare.7476359).

Pseudoknot Structures

The online tool DotKnot was used to predict pseudoknots in certain viral genome sequences [124,125]. PseudoViewer 3.0 was used for visualization of the predicted pseudoknots [126].

Read distribution plots

Distribution of reads along the viral genome segment(s) was visualized by first aligning the reads to the viral genome using Bowtie2, then calculating the hits per nucleotide along the viral sequence(s) using the custom Perl script plotRegion_RNAseq_hitsPerNT.pl. The resulting output table was the input for an R script to create hits per nucleotide heat maps for each viral segment. The R code for each plot, along with the input data table, is available through the FigShare FileSet (DOI: 10.6084/m9.figshare.7476359).

Phylogenetic analysis

To identify related sequences for phylogenetic analysis, the predicted amino acid sequence for the RdRP was used to query the 'nr' database at NCBI. Protein sequences of the top hits, generally with e-values approaching or equal to 0.0, were downloaded. Partial sequences were excluded. Known viral sequences that appeared as the top hit for multiple viruses identified in

this study were deduplicated. The complete list of Accession IDs used for phylogenetic analysis is available in Supp. Table S2.

To build the trees, amino acid sequences were aligned with MAFFT [127], and resultant alignments were degapped and trimmed using the tool trimAL [128] with a gap threshold of 0.9 (allowing for a gap in up to 10% of the sequences at a given position). Maximum likelihood phylogenetic trees were constructed with RAxML version 8.2.4 [129] using the LG model for amino acid substitution. RAxML tree files, with branch lengths, were viewed in Dendroscope version 3.5.7 [130] and exported as PDF for figure preparation.

Percent Identity Matrices

Percent identities among protein coding sequences, either full length or motif-specific, were determined via multiple sequence alignment using Clustal Omega [131]. In instances where two different proteins were evaluated and then combined into one matrix, first the RdRP protein sequences were analyzed; the order of the output in the percent identity matrix was then applied to the second set of protein sequences to be analyzed and the option “Order: Input” was selected in Clustal-Omega. Percent identity matrices were converted to heat map plots using a custom R script. This script along with the input matrices are available through the FigShare FileSet (DOI: 10.6084/m9.figshare.7476359).

Accession numbers

Nucleotide sequences of the mycoviral genomes identified and analyzed in this study were deposited at NCBI; GenBank accession numbers are listed in Table 1.

Acknowledgments

We are deeply grateful to researchers around the world that have shared their transcriptomic data, both before and after publication. We particularly wish to acknowledge the sequencing data produced by various groups at the US Department of Energy Joint Genome Institute (<http://www.jgi.doe.gov>). We also thank the Bioinformatics Facility staff at the Danforth Center for excellent computational support.

Author Contributions

KBG, EEH, and JCC designed the research. KBG, EEH, and RLA performed the research and analyzed the data. KBG, EEH, and JCC drafted the paper; all authors edited and approved the manuscript.

References

1. Bekal S, Domier LL, Niblack TL, Lambert KN. Discovery and initial analysis of novel viral genomes in the soybean cyst nematode. *J Gen Virol*. 2011;92: 1870–1879.
doi:10.1099/vir.0.030585-0
2. Mushegian A, Shipunov A, Elena SF. Changes in the composition of the RNA virome mark evolutionary transitions in green plants. *BMC Biol*. 2016;14: 68.
doi:10.1186/s12915-016-0288-8
3. Debat HJ. An RNA Virome Associated to the Golden Orb-Weaver Spider *Nephila clavipes*. *Front Microbiol*. 2017;8: 2097. doi:10.3389/fmicb.2017.02097
4. Nerva L, Ciuffo M, Vallino M, Margaria P, Varese GC, Gnani G, et al. Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi. *Virus Res*. 2016;219: 22–38. doi:10.1016/j.virusres.2015.10.028
5. Marzano S-YL, Nelson BD, Ajayi-Oyetunde O, Bradley CA, Hughes TJ, Hartman GL, et al. Identification of Diverse Mycoviruses through Metatranscriptomics Characterization of the Viromes of Five Major Fungal Plant Pathogens. *J Virol*. 2016;90: 6846–6863.
doi:10.1128/JVI.00357-16
6. Marzano S-YL, Domier LL. Novel mycoviruses discovered from metatranscriptomics survey of soybean phyllosphere phytobiomes. *Virus Res*. 2016;213: 332–342.
doi:10.1016/j.virusres.2015.11.002
7. Hansen JL, Long AM, Schultz SC. Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure*. 1997;5: 1109–1122. doi:10.1016/S0969-2126(97)00261-X
8. Campo S, Gilbert KB, Carrington JC. Small RNA-Based Antiviral Defense in the Phytopathogenic Fungus *Colletotrichum higginsianum*. *PLoS Pathog*. 2016;12: e1005640.
doi:10.1371/journal.ppat.1005640
9. Valero-Jiménez CA, Faino L, Spring In't Veld D, Smit S, Zwaan BJ, van Kan JAL. Comparative genomics of *Beauveria bassiana*: uncovering signatures of virulence against

- mosquitoes. *BMC Genomics*. 2016;17: 986. doi:10.1186/s12864-016-3339-1
10. Zhu HJ, Chen D, Zhong J, Zhang SY, Gao BD. A novel mycovirus identified from the rice false smut fungus *Ustilagoidea virescens*. *Virus Genes*. 2015;51: 159–162. doi:10.1007/s11262-015-1212-y
 11. Koloniuk I, Hrabáková L, Petrzik K. Molecular characterization of a novel amalgavirus from the entomopathogenic fungus *Beauveria bassiana*. *Arch Virol*. 2015;160: 1585–1588. doi:10.1007/s00705-015-2416-0
 12. Kotta-Loizou I, Sipkova J, Coutts RHA. Identification and sequence determination of a novel double-stranded RNA mycovirus from the entomopathogenic fungus *Beauveria bassiana*. *Arch Virol*. Springer Vienna; 2015;160: 873–875. doi:10.1007/s00705-014-2332-8
 13. Lin Y, Zhang H, Zhao C, Liu S, Guo L. The complete genome sequence of a novel mycovirus from *Alternaria longipes* strain HN28. *Arch Virol*. 2015;160: 577–580. doi:10.1007/s00705-014-2218-9
 14. Aoki N, Moriyama H, Kodama M, Arie T, Teraoka T, Fukuhara T. A novel mycovirus associated with four double-stranded RNAs affects host fungal growth in *Alternaria alternata*. *Virus Res*. 2009;140: 179–187. doi:10.1016/j.virusres.2008.12.003
 15. Kozlakidis Z, Herrero N, Ozkan S, Kanhayuwa L, Jamal A, Bhatti MF, et al. Sequence determination of a quadripartite dsRNA virus isolated from *Aspergillus foetidus*. *Arch Virol*. 2013;158: 267–272. doi:10.1007/s00705-012-1362-3
 16. Sinha S, Flibotte S, Neira M, Formby S, Plemenitaš A, Cimerman NG, et al. Insight into the Recent Genome Duplication of the Halophilic Yeast *Hortaea werneckii*: Combining an Improved Genome with Gene Expression and Chromatin Structure. *G3*. 2017;7: 2015–2022. doi:10.1534/g3.117.040691
 17. King AMQ, Lefkowitz E, Adams MJ, Carstens EB. *Virus Taxonomy: Ninth Report of the International Committee on Taxonomy of Viruses* [Internet]. Elsevier; 2011. Available: <https://market.android.com/details?id=book-aFYaE9KXEXUC>
 18. Baeza M, Bravo N, Sanhueza M, Flores O, Villarreal P, Cifuentes V. Molecular

- characterization of totiviruses in *Xanthophyllomyces dendrorhous*. *Virology*. 2012;9: 140.
doi:10.1186/1743-422X-9-140
19. Shi M, Lin X-D, Tian J-H, Chen L-J, Chen X, Li C-X, et al. Redefining the invertebrate RNA virosphere. *Nature*. 2016; doi:10.1038/nature20167
 20. International Committee on Taxonomy of Viruses, King AMQ. *Virus Taxonomy: Classification and Nomenclature of Viruses : Ninth Report of the International Committee on Taxonomy of Viruses* [Internet]. Elsevier; 2012. Available: <https://market.android.com/details?id=book-KXRCYay3pH4C>
 21. Ghabrial SA, Nibert ML. Victorivirus, a new genus of fungal viruses in the family Totiviridae. *Arch Virol*. 2009;154: 373–379. doi:10.1007/s00705-008-0272-x
 22. Wang M, Sun X, Zhu C, Xu Q, Ruan R, Yu D, et al. PdbrlA, PdabaA and PdwetA control distinct stages of conidiogenesis in *Penicillium digitatum*. *Res Microbiol*. 2015;166: 56–65. doi:10.1016/j.resmic.2014.12.003
 23. Quandt CA, Di Y, Elser J, Jaiswal P, Spatafora JW. Differential Expression of Genes Involved in Host Recognition, Attachment, and Degradation in the Mycoparasite *Tolyposcladium ophioglossoides*. *G3*. 2016;6: 731–741. doi:10.1534/g3.116.027045
 24. Komatsu K, Katayama Y, Omatsu T, Mizutani T, Fukuhara T, Kodama M, et al. Genome sequence of a novel victorivirus identified in the phytopathogenic fungus *Alternaria arborescens*. *Arch Virol*. 2016;161: 1701–1704. doi:10.1007/s00705-016-2796-9
 25. Romo M, Leuchtman A, García B, Zabalgogezcoa I. A totivirus infecting the mutualistic fungal endophyte *Epichloë festucae*. *Virus Res*. 2007;124: 38–43. doi:10.1016/j.virusres.2006.09.008
 26. Huang S, Ghabrial SA. Organization and expression of the double-stranded RNA genome of *Helminthosporium victoriae* 190S virus, a totivirus infecting a plant pathogenic filamentous fungus. *Proc Natl Acad Sci U S A*. 1996;93: 12541–12546. Available: <https://www.ncbi.nlm.nih.gov/pubmed/8901618>
 27. Li H, Havens WM, Nibert ML, Ghabrial SA. RNA sequence determinants of a coupled termination-reinitiation strategy for downstream open reading frame translation in

- Helminthosporium victoriae virus 190S and other victoriviruses (Family Totiviridae). *J Virol.* 2011;85: 7343–7352. doi:10.1128/JVI.00364-11
28. Preisig O, Wingfield BD, Wingfield MJ. Coinfection of a fungal pathogen by two distinct double-stranded RNA viruses. *Virology.* 1998;252: 399–406. doi:10.1006/viro.1998.9480
 29. Niu Y, Zhang T, Zhu Y, Yuan Y, Wang S, Liu J, et al. Isolation and characterization of a novel mycovirus from *Penicillium digitatum*. *Virology.* 2016;494: 15–22. doi:10.1016/j.virol.2016.04.004
 30. Vainio EJ, Chiba S, Ghabrial SA, Maiss E, Roossinck M, Sabanadzovic S, et al. ICTV Virus Taxonomy Profile: Partitiviridae. *J Gen Virol.* 2018;99: 17–18. doi:10.1099/jgv.0.000985
 31. Mondo SJ, Dannebaum RO, Kuo RC, Louie KB, Bewick AJ, LaButti K, et al. Widespread adenine N6-methylation of active genes in fungi. *Nat Genet.* 2017;49: 964–968. doi:10.1038/ng.3859
 32. Wang Y, Lim L, DiGuistini S, Robertson G, Bohlmann J, Breuil C. A specialized ABC efflux transporter GcABC-G1 confers monoterpene resistance to *Grosmannia clavigera*, a bark beetle-associated fungal pathogen of pine trees. *New Phytol.* 2013;197: 886–898. doi:10.1111/nph.12063
 33. Liu K, Zhang W, Lai Y, Xiang M, Wang X, Zhang X, et al. *Drechlerella stenobrocha* genome illustrates the mechanism of constricting rings and the origin of nematode predation in fungi. *BMC Genomics.* 2014;15: 114. doi:10.1186/1471-2164-15-114
 34. Yang L, Xie L, Xue B, Goodwin PH, Quan X, Zheng C, et al. Comparative transcriptome profiling of the early infection of wheat roots by *Gaeumannomyces graminis* var. *tritici*. *PLoS One.* 2015;10: e0120691. doi:10.1371/journal.pone.0120691
 35. Covelli L, Kozlakidis Z, Di Serio F, Citir A, Açikgöz S, Hernández C, et al. Sequences of the smallest double-stranded RNAs associated with cherry chlorotic rusty spot and Amasya cherry diseases. *Arch Virol.* 2008;153: 759–762. doi:10.1007/s00705-008-0039-4
 36. Lin H, Kazlauskas RJ, Travisano M. Developmental evolution facilitates rapid adaptation.

Sci Rep. 2017;7: 15891. doi:10.1038/s41598-017-16229-0

37. Vanheule A, Audenaert K, Warris S, van de Geest H, Schijlen E, Höfte M, et al. Living apart together: crosstalk between the core and supernumerary genomes in a fungal plant pathogen. *BMC Genomics*. 2016;17: 670. doi:10.1186/s12864-016-2941-6
38. Osaki H, Sasaki A, Nomiya K, Tomioka K. Multiple virus infection in a single strain of *Fusarium poae* shown by deep sequencing. *Virus Genes*. 2016;52: 835–847. doi:10.1007/s11262-016-1379-x
39. Compel P, Papp I, Bibó M, Fekete C, Hornok L. Genetic interrelationships and genome organization of double-stranded RNA elements of *Fusarium poae*. *Virus Genes*. 1999;18: 49–56. Available: <https://www.ncbi.nlm.nih.gov/pubmed/10334037>
40. Luo J, Qiu H, Cai G, Wagner NE, Bhattacharya D, Zhang N. Phylogenomic analysis uncovers the evolutionary history of nutrition and infection mode in rice blast fungus and other Magnaporthales. *Sci Rep*. 2015;5: 9448. doi:10.1038/srep09448
41. Nerva L, Silvestri A, Ciuffo M, Palmano S, Varese GC, Turina M. Transmission of *Penicillium aurantiogriseum* partiti-like virus 1 to a new fungal host (*Cryphonectria parasitica*) confers higher resistance to salinity and reveals adaptive genomic changes. *Environ Microbiol*. 2017;19: 4480–4492. doi:10.1111/1462-2920.13894
42. Horn F, Linde J, Mattern DJ, Walther G, Guthke R, Brakhage AA, et al. Draft Genome Sequence of the Fungus *Penicillium brasilianum* MG11. *Genome Announc*. 2015;3. doi:10.1128/genomeA.00724-15
43. Xavier A da S, Barros APO de, Godinho MT, Zerbini FM, Souza F de O, Bruckner FP, et al. A novel mycovirus associated to *Alternaria alternata* comprises a distinct lineage in Partitiviridae. *Virus Res*. 2017;244: 21–26. doi:10.1016/j.virusres.2017.10.007
44. Wang L, Jiang J, Wang Y, Hong N, Zhang F, Xu W, et al. Hypovirulence of the phytopathogenic fungus *Botryosphaeria dothidea*: association with a coinfecting chrysovirus and a partitivirus. *J Virol*. 2014;88: 7517–7527. doi:10.1128/JVI.00538-14
45. Ghabrial SA, Castón JR, Coutts RHA, Hillman BI, Jiang D, Kim D-H, et al. ICTV Virus

- Taxonomy Profile: Chrysoviridae. *J Gen Virol.* 2018;99: 19–20. doi:10.1099/jgv.0.000994
46. Wang J-J, Bai W-W, Zhou W, Liu J, Chen J, Liu X-Y, et al. Transcriptomic analysis of two *Beauveria bassiana* strains grown on cuticle extracts of the silkworm uncovers their different metabolic response at early infection stage. *J Invertebr Pathol.* 2017;145: 45–54. doi:10.1016/j.jip.2017.03.010
 47. Herrero N. Identification and sequence determination of a new chrysovirus infecting the entomopathogenic fungus *Isaria javanica*. *Arch Virol.* 2017;162: 1113–1117. doi:10.1007/s00705-016-3194-z
 48. Jiang D, Ghabrial SA. Molecular characterization of *Penicillium chrysogenum* virus: reconsideration of the taxonomy of the genus Chrysovirus. *J Gen Virol.* 2004;85: 2111–2121. doi:10.1099/vir.0.79842-0
 49. Covelli L, Coutts RHA, Di Serio F, Citir A, Açikgöz S, Hernández C, et al. Cherry chlorotic rusty spot and Amasya cherry diseases are associated with a complex pattern of mycoviral-like double-stranded RNAs. I. Characterization of a new species in the genus Chrysovirus. *J Gen Virol.* 2004;85: 3389–3397. doi:10.1099/vir.0.80181-0
 50. Kim J-M, Park J-A, Park S-M, Cha B-J, Yang M-S, Kim D-H. Nucleotide sequences of four segments of chrysovirus in Korean *Cryphonectria nitschkei* BS122 strain. *Virus Genes.* 2010;41: 292–294. doi:10.1007/s11262-010-0495-2
 51. Khalifa ME, Pearson MN. Molecular characterisation of an endornavirus infecting the phytopathogen *Sclerotinia sclerotiorum*. *Virus Res.* 2014;189: 303–309. doi:10.1016/j.virusres.2014.06.010
 52. Moriyama H, Horiuchi H, Koga R, Fukuhara T. Molecular characterization of two endogenous double-stranded RNAs in rice and their inheritance by interspecific hybrids. *J Biol Chem.* 1999;274: 6882–6888. doi:10.1074/jbc.274.11.6882
 53. Ong JWL, Li H, Sivasithamparam K, Dixon KW, Jones MGK, Wylie SJ. Novel Endorna-like viruses, including three with two open reading frames, challenge the membership criteria and taxonomy of the Endornaviridae. *Virology.* 2016;499: 203–211. doi:10.1016/j.virol.2016.08.019

54. Das S, Falloon RE, Stewart A, Pitman AR. Molecular characterisation of an endornavirus from *Rhizoctonia solani* AG-3PT infecting potato. *Fungal Biol.* 2014;118: 924–934. doi:10.1016/j.funbio.2014.08.003
55. Moriyama H, Nitta T, Fukuhara T. Double-stranded RNA in rice: a novel RNA replicon in plants. *Mol Gen Genet.* 1995;248: 364–369. Available: <https://www.ncbi.nlm.nih.gov/pubmed/7565598>
56. Pandey B, Naidu RA, Grove GG. Next generation sequencing analysis of double-stranded RNAs from sweet cherry powdery mildew fungus *Podosphaera prunicola*. *J Plant Pathol.* 2018; doi:10.1007/s42161-018-0092-0
57. Adams MJ, Adkins S, Bragard C, Gilmer D, Li D, MacFarlane SA, et al. ICTV Virus Taxonomy Profile: Virgaviridae. *J Gen Virol.* 2017;98: 1999–2000. doi:10.1099/jgv.0.000884
58. Mascia T, Nigro F, Abdallah A, Ferrara M, De Stradis A, Faedda R, et al. Gene silencing and gene expression in phytopathogenic fungi using a plant virus vector. *Proc Natl Acad Sci U S A.* 2014;111: 4291–4296. doi:10.1073/pnas.1315668111
59. Andika IB, Wei S, Cao C, Salaipeth L, Kondo H, Sun L. Phytopathogenic fungus hosts a plant virus: A naturally occurring cross-kingdom viral infection. *Proc Natl Acad Sci U S A.* 2017;114: 12267–12272. doi:10.1073/pnas.1714916114
60. Mosier AC, Miller CS, Frischkorn KR, Ohm RA, Li Z, LaButti K, et al. Fungi Contribute Critical but Spatially Varying Roles in Nitrogen and Carbon Cycling in Acid Mine Drainage. *Front Microbiol.* 2016;7: 238. doi:10.3389/fmicb.2016.00238
61. Deneff VJ, Mueller RS, Banfield JF. AMD biofilms: using model communities to study microbial evolution and ecological complexity in nature. *ISME J.* 2010;4: 599–610. doi:10.1038/ismej.2009.158
62. Andersson AF, Banfield JF. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science.* 2008;320: 1047–1050. doi:10.1126/science.1157358
63. Steindorff AS, Ramada MHS, Coelho ASG, Miller RNG, Pappas GJ Jr, Ulhoa CJ, et al.

- Identification of mycoparasitism-related genes against the phytopathogen *Sclerotinia sclerotiorum* through transcriptome and expression profile analysis in *Trichoderma harzianum*. *BMC Genomics*. 2014;15: 204. doi:10.1186/1471-2164-15-204
64. Knapp DG, Németh JB, Barry K, Hainaut M, Henrissat B, Johnson J, et al. Comparative genomics provides insights into the lifestyle and reveals functional heterogeneity of dark septate endophytic fungi. *Sci Rep*. 2018;8: 6321. doi:10.1038/s41598-018-24686-4
 65. Cañizares MC, López-Escudero FJ, Pérez-Artés E, García-Pedrajas MD. Characterization of a novel single-stranded RNA mycovirus related to invertebrate viruses from the plant pathogen *Verticillium dahliae*. *Arch Virol*. 2018;163: 771–776. doi:10.1007/s00705-017-3644-2
 66. Ai Y-P, Zhong J, Chen C-Y, Zhu H-J, Gao B-D. A novel single-stranded RNA virus isolated from the rice-pathogenic fungus *Magnaporthe oryzae* with similarity to members of the family Tombusviridae. *Arch Virol*. 2016;161: 725–729. doi:10.1007/s00705-015-2683-9
 67. Preisig O, Moleleki N, Smit WA, Wingfield BD, Wingfield MJ. A novel RNA mycovirus in a hypovirulent isolate of the plant pathogen *Diaporthe ambigua*. *J Gen Virol*. 2000;81: 3107–3114. doi:10.1099/0022-1317-81-12-3107
 68. Poch O, Sauvaget I, Delarue M, Tordo N. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J*. 1989;8: 3867–3874. Available: <https://www.ncbi.nlm.nih.gov/pubmed/2555175>
 69. Vázquez AL, Alonso JM, Parra F. Mutation analysis of the GDD sequence motif of a calicivirus RNA-dependent RNA polymerase. *J Virol*. 2000;74: 3888–3891. Available: <https://www.ncbi.nlm.nih.gov/pubmed/10729164>
 70. Lohmann V, Körner F, Herian U, Bartenschlager R. Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J Virol*. 1997;71: 8416–8428. Available: <https://www.ncbi.nlm.nih.gov/pubmed/9343198>
 71. Jablonski SA, Morrow CD. Mutation of the aspartic acid residues of the GDD sequence motif of poliovirus RNA-dependent RNA polymerase results in enzymes with altered metal

- ion requirements for activity. *J Virol.* 1995;69: 1532–1539. Available: <https://www.ncbi.nlm.nih.gov/pubmed/7853486>
72. O'Reilly EK, Kao CC. Analysis of RNA-dependent RNA polymerase structure and function as guided by known polymerase structures and computer predictions of secondary structure. *Virology.* 1998;252: 287–303. doi:10.1006/viro.1998.9463
 73. Hisano S, Zhang R, Faruk MI, Kondo H, Suzuki N. A neo-virus lifestyle exhibited by a (+)ssRNA virus hosted in an unrelated dsRNA virus: Taxonomic and evolutionary considerations. *Virus Res.* 2018;244: 75–83. doi:10.1016/j.virusres.2017.11.006
 74. Kozlakidis Z, Herrero N, Ozkan S, Bhatti MF, Coutts RHA. A novel dsRNA element isolated from the *Aspergillus foetidus* mycovirus complex. *Arch Virol.* 2013;158: 2625–2628. doi:10.1007/s00705-013-1779-3
 75. Zhang R, Hisano S, Tani A, Kondo H, Kanematsu S, Suzuki N. A capsidless ssRNA virus hosted by an unrelated dsRNA virus. *Nat Microbiol.* 2016;1: 15001. doi:10.1038/nmicrobiol.2015.1
 76. Koonin EV, Wolf YI, Nagasaki K, Dolja VV. The Big Bang of picorna-like virus evolution antedates the radiation of eukaryotic supergroups. *Nat Rev Microbiol.* 2008;6: 925–939. doi:10.1038/nrmicro2030
 77. Kotta-Loizou I, Coutts RHA. Mycoviruses in Aspergilli: A Comprehensive Review. *Front Microbiol.* 2017;8: 1699. doi:10.3389/fmicb.2017.01699
 78. Hillman BI, Cai G. The family narnaviridae: simplest of RNA viruses. *Adv Virus Res.* 2013;86: 149–176. doi:10.1016/B978-0-12-394315-6.00006-4
 79. Osawa S, Jukes TH, Watanabe K, Muto A. Recent evidence for evolution of the genetic code. *Microbiol Rev.* 1992;56: 229–264. Available: <https://www.ncbi.nlm.nih.gov/pubmed/1579111>
 80. Ashwin NMR, Barnabas L, Ramesh Sundar A, Malathi P, Viswanathan R, Masi A, et al. Comparative secretome analysis of *Colletotrichum falcatum* identifies a cerato-platanin protein (EPL1) as a potential pathogen-associated molecular pattern (PAMP) inducing systemic resistance in sugarcane. *J Proteomics.* 2017;169: 2–20.

doi:10.1016/j.jprot.2017.05.020

81. Li Y, Hsiang T, Yang R-H, Hu X-D, Wang K, Wang W-J, et al. Comparison of different sequencing and assembly strategies for a repeat-rich fungal genome, *Ophiocordyceps sinensis*. *J Microbiol Methods*. 2016;128: 1–6. doi:10.1016/j.mimet.2016.06.025
82. Turina M, Hillman BI, Izadpanah K, Rastgou M, Rosa C, Ictv Report Consortium. ICTV Virus Taxonomy Profile: Ourmiavirus. *J Gen Virol*. 2017;98: 129–130. doi:10.1099/jgv.0.000725
83. Hrabáková L, Koloniuk I, Petrzik K. Phomopsis longicolla RNA virus 1 - Novel virus at the edge of myco- and plant viruses. *Virology*. 2017;506: 14–18. doi:10.1016/j.virol.2017.03.003
84. Donaire L, Rozas J, Ayllón MA. Molecular characterization of Botrytis ourmia-like virus, a mycovirus close to the plant pathogenic genus Ourmiavirus. *Virology*. 2016;489: 158–164. doi:10.1016/j.virol.2015.11.027
85. Suzuki N, Ghabrial SA, Kim K-H, Pearson M, Marzano S-YL, Yaegashi H, et al. ICTV Virus Taxonomy Profile: Hypoviridae. *J Gen Virol*. 2018;99: 615–616. doi:10.1099/jgv.0.001055
86. Fahima T, Kazmierczak P, Hansen DR, Pfeiffer P, Van Alfen NK. Membrane-associated replication of an unencapsidated double-strand RNA of the fungus, *Cryphonectria parasitica*. *Virology*. 1993;195: 81–89. doi:10.1006/viro.1993.1348
87. Jacob-Wilk D, Turina M, Van Alfen NK. Mycovirus cryphonectria hypovirus 1 elements cofractionate with trans-Golgi network membranes of the fungal host *Cryphonectria parasitica*. *J Virol*. 2006;80: 6588–6596. doi:10.1128/JVI.02519-05
88. Hulvey J, Popko JT Jr, Sang H, Berg A, Jung G. Overexpression of ShCYP51B and ShatrD in *Sclerotinia homoeocarpa* isolates exhibiting practical field resistance to a demethylation inhibitor fungicide. *Appl Environ Microbiol*. 2012;78: 6674–6682. doi:10.1128/AEM.00417-12
89. Bech L, Busk PK, Lange L, Others. Cell wall degrading enzymes in *Trichoderma asperellum* grown on wheat bran. *Fungal Genom Biol*. 2015;4: 116. Available:

<http://dx.doi.org/>

90. Wang S, Zhang J, Li P, Qiu D, Guo L. Transcriptome-Based Discovery of *Fusarium graminearum* Stress Responses to FgHV1 Infection. *Int J Mol Sci.* 2016;17.
doi:10.3390/ijms17111922
91. Li P, Zhang H, Chen X, Qiu D, Guo L. Molecular characterization of a novel hypovirus from the plant pathogenic fungus *Fusarium graminearum*. *Virology.* 2015;481: 151–160.
doi:10.1016/j.virol.2015.02.047
92. Yaegashi H, Kanematsu S, Ito T. Molecular characterization of a new hypovirus infecting a phytopathogenic fungus, *Valsa ceratosperma*. *Virus Res.* 2012;165: 143–150.
doi:10.1016/j.virusres.2012.02.008
93. Shapira R, Choi GH, Nuss DL. Virus-like genetic organization and expression strategy for a double-stranded RNA genetic element associated with biological control of chestnut blight. *EMBO J.* 1991;10: 731–739. Available:
<https://www.ncbi.nlm.nih.gov/pubmed/2009854>
94. Hillman BI, Halpern BT, Brown MP. A viral dsRNA element of the chestnut blight fungus with a distinct genetic organization. *Virology.* 1994;201: 241–250.
doi:10.1006/viro.1994.1289
95. Arjona-Lopez JM, Telengech P, Jamal A, Hisano S, Kondo H, Yelin MD, et al. Novel, diverse RNA viruses from Mediterranean isolates of the phytopathogenic fungus, *Rosellinia necatrix*: insights into evolutionary biology of fungal viruses. *Environ Microbiol.* 2018;20: 1464–1483. doi:10.1111/1462-2920.14065
96. Li P, Chen X, He H, Qiu D, Guo L. Complete Genome Sequence of a Novel Hypovirus from the Phytopathogenic Fungus *Fusarium langsethiae*. *Genome Announc.* 2017;5.
doi:10.1128/genomeA.01722-16
97. Lehr NA, Wang Z, Li N, Hewitt DA, López-Giráldez F, Trail F, et al. Gene expression differences among three *Neurospora* species reveal genes required for sexual reproduction in *Neurospora crassa*. *PLoS One.* 2014;9: e110398.
doi:10.1371/journal.pone.0110398

98. Kellner R, Bhattacharyya A, Poppe S, Hsu TY, Brem RB, Stukenbrock EH. Expression profiling of the wheat pathogen *Zymoseptoria tritici* reveals genomic patterns of transcription and host-specific regulatory programs. *Genome Biol Evol.* 2014;6: 1353–1365. doi:10.1093/gbe/evu101
99. Rudd JJ, Kanyuka K, Hassani-Pak K, Derbyshire M, Andongabo A, Devonshire J, et al. Transcriptome and metabolite profiling of the infection cycle of *Zymoseptoria tritici* on wheat reveals a biphasic interaction with plant immunity involving differential pathogen chromosomal contributions and a variation on the hemibiotrophic lifestyle definition. *Plant Physiol.* 2015;167: 1158–1185. doi:10.1104/pp.114.255927
100. Omrane S, Sghyer H, Audéon C, Lanen C, Duplaix C, Walker A-S, et al. Fungicide efflux and the MgMFS1 transporter contribute to the multidrug resistance phenotype in *Zymoseptoria tritici* field isolates. *Environ Microbiol.* 2015;17: 2805–2823. doi:10.1111/1462-2920.12781
101. Zhang R, Liu S, Chiba S, Kondo H, Kanematsu S, Suzuki N. A novel single-stranded RNA virus isolated from a phytopathogenic filamentous fungus, *Rosellinia necatrix*, with similarity to hypo-like viruses. *Front Microbiol.* 2014;5: 360. doi:10.3389/fmicb.2014.00360
102. Hrabáková L, Grum-Grzhimaylo AA, Koloniuk I, Debets AJM, Sarkisova T, Petrzik K. The alkalophilic fungus *Sodiomyces alkalinus* hosts beta- and gammapartitiviruses together with a new fusarivirus. *PLoS One.* 2017;12: e0187799. doi:10.1371/journal.pone.0187799
103. Liu R, Cheng J, Fu Y, Jiang D, Xie J. Molecular Characterization of a Novel Positive-Sense, Single-Stranded RNA Mycovirus Infecting the Plant Pathogenic Fungus *Sclerotinia sclerotiorum*. *Viruses.* 2015;7: 2470–2484. doi:10.3390/v7052470
104. Liu H, Fu Y, Jiang D, Li G, Xie J, Peng Y, et al. A novel mycovirus that is related to the human pathogen hepatitis E virus and rubi-like viruses. *J Virol.* 2009;83: 1981–1991. doi:10.1128/JVI.01897-08
105. Bartholomäus A, Wibberg D, Winkler A, Pühler A, Schlüter A, Varrelmann M. Deep Sequencing Analysis Reveals the Mycoviral Diversity of the Virome of an Avirulent Isolate of *Rhizoctonia solani* AG-2-2 IV. *PLoS One.* 2016;11: e0165965.

doi:10.1371/journal.pone.0165965

106. Koonin EV, Dolja VV, Krupovic M. Origins and evolution of viruses of eukaryotes: The ultimate modularity. *Virology*. 2015;479-480: 2–25. doi:10.1016/j.virol.2015.02.039
107. Hacquard S, Kracher B, Hiruma K, Münch PC, Garrido-Oter R, Thon MR, et al. Survival trade-offs in plant roots during colonization by closely related beneficial and pathogenic fungi. *Nat Commun*. 2016;7: 11362. doi:10.1038/ncomms11362
108. Grandaubert J, Bhattacharyya A, Stukenbrock EH. RNA-seq-Based Gene Annotation and Comparative Genomics of Four Fungal Grass Pathogens in the Genus *Zymoseptoria* Identify Novel Orphan Genes and Species-Specific Invasions of Transposable Elements. *G3* . 2015;5: 1323–1333. doi:10.1534/g3.115.017731
109. Rajakaruna P, Khandekar S, Meulia T, Leisner SM. Identification and Host Relations of Turnip ringspot virus, A Novel Comovirus from Ohio. *Plant Dis. Scientific Societies*; 2007;91: 1212–1220. doi:10.1094/PDIS-91-10-1212
110. Roossinck MJ. The good viruses: viral mutualistic symbioses. *Nat Rev Microbiol*. 2011;9: 99–108. doi:10.1038/nrmicro2491
111. Schmitt MJ, Breinig F. The viral killer system in yeast: from molecular biology to application. *FEMS Microbiol Rev*. 2002;26: 257–276. Available: <https://www.ncbi.nlm.nih.gov/pubmed/12165427>
112. Márquez LM, Redman RS, Rodriguez RJ, Roossinck MJ. A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science*. 2007;315: 513–515. doi:10.1126/science.1136237
113. Niu Y, Yuan Y, Mao J, Yang Z, Cao Q, Zhang T, et al. Characterization of two novel mycoviruses from *Penicillium digitatum* and the related fungicide resistance analysis. *Sci Rep*. 2018;8: 5513. doi:10.1038/s41598-018-23807-3
114. Chiba S, Suzuki N. Highly activated RNA silencing via strong induction of dicer by one virus can interfere with the replication of an unrelated virus. *Proc Natl Acad Sci U S A*. 2015;112: E4911–8. doi:10.1073/pnas.1509151112

115. Segers GC, Zhang X, Deng F, Sun Q, Nuss DL. Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. *Proc Natl Acad Sci U S A*. 2007;104: 12902–12906. doi:10.1073/pnas.0702500104
116. Mochama P, Jadhav P, Neupane A, Lee Marzano S-Y. Mycoviruses as Triggers and Targets of RNA Silencing in White Mold Fungus *Sclerotinia sclerotiorum*. *Viruses*. 2018;10. doi:10.3390/v10040214
117. Kirk, PM, Cannon, PF, David, JC, Stalpers, JA, editor. *Ainsworth and Bisby's Dictionary of Fungi*. CABI Bioscience; 2001.
118. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9: 357–359. doi:10.1038/nmeth.1923
119. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc*. 2013;8: 1494–1512. doi:10.1038/nprot.2013.084
120. TransDecoder WebCitation [Internet]. [cited 16 Aug 2017]. Available: transdecoder.github.io
121. HMMER [Internet]. [cited 14 Jun 2018]. Available: <http://hmmer.org>
122. Home - ORFfinder - NCBI [Internet]. [cited 20 Dec 2017]. Available: <https://www.ncbi.nlm.nih.gov/orffinder/>
123. SankeyMATIC (BETA): A Sankey diagram builder for everyone [Internet]. [cited 14 Jun 2018]. Available: <http://sankeymatic.com/>
124. Sperschneider J, Datta A. DotKnot: pseudoknot prediction using the probability dot plot under a refined energy model. *Nucleic Acids Res*. 2010;38: e103. doi:10.1093/nar/gkq021
125. Sperschneider J, Datta A, Wise MJ. Heuristic RNA pseudoknot prediction including intramolecular kissing hairpins. *RNA*. 2011;17: 27–38. doi:10.1261/rna.2394511
126. Byun Y, Han K. PseudoViewer3: generating planar drawings of large-scale RNA structures with pseudoknots. *Bioinformatics*. 2009;25: 1435–1437. doi:10.1093/bioinformatics/btp252

127. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30: 772–780. doi:10.1093/molbev/mst010
128. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics.* 2009;25: 1972–1973. doi:10.1093/bioinformatics/btp348
129. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 2006;22: 2688–2690. doi:10.1093/bioinformatics/btl446
130. Huson DH, Scornavacca C. Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst Biol.* 2012;61: 1061–1067. doi:10.1093/sysbio/sys062
131. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S, et al. The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res.* 2015;43: W580–4. doi:10.1093/nar/gkv279

Table S1: Complete list of BioProjects and SRA datasets analyzed

Genus	species	BioProject	SRAFile(s)	Result
<i>Acidomyces</i>	<i>richmondensis</i>	PRJNA250470	SRR1797521	Virus
<i>Aspergillus</i>	<i>ellipticus</i>	PRJNA250911	SRR1799565	Virus
<i>Aspergillus</i>	<i>heteromorphus</i>	PRJNA250969	SRR1799577	Virus
<i>Aspergillus</i>	<i>homomorphus</i>	PRJNA250984	SRR1799578;SRR3501509	Virus
<i>Aspergillus</i>	<i>neoniger</i>	PRJNA250996	SRR1801411	Virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA260878	SRR3269778	Virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA306902	SRR3043102	Virus
<i>Clohesyomyces</i>	<i>aquaticus</i>	PRJNA372822	SRR5494002	Virus
<i>Colletotrichum</i>	<i>caudatum</i>	PRJNA262373	SRR5145569	Virus
<i>Colletotrichum</i>	<i>eremochloae</i>	PRJNA262412	SRR5166050	Virus
<i>Colletotrichum</i>	<i>falcatum</i>	PRJNA272832	SRR1765657	Virus
<i>Colletotrichum</i>	<i>higginsianum</i>	PRJNA264846	SRR1630207	Virus
<i>Colletotrichum</i>	<i>higginsianum</i>	PRJNA148307	SRR364395	Virus
<i>Colletotrichum</i>	<i>navitas</i>	PRJNA262225	SRR5166338	Virus
<i>Colletotrichum</i>	<i>zoysiae</i>	PRJNA262216	SRR5166062	Virus
<i>Cryphonectria</i>	<i>parasitica</i>	PRJNA369604	SRR5235483	Virus
<i>Delitschia</i>	<i>confertaspera</i>	PRJNA250740	SRR3440303	Virus
<i>Drechslerella</i>	<i>stenobrocha</i>	PRJNA236481	SRR1145648	Virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA263651	SRR4445678	Virus
<i>Fusarium</i>	<i>poae</i>	PRJNA319914	SRR3953125	Virus
<i>Gaeumannomyces</i>	<i>tritici</i>	PRJNA268052	SRR1664730	Virus
<i>Grossmannia</i>	<i>clavigera</i>	PRJNA184372	SRR636711	Virus
<i>Gyromitra</i>	<i>esculenta</i>	PRJNA372840	SRR5491178	Virus
<i>Hortaea</i>	<i>wemeckii</i>	PRJNA356640	SRR5086622	Virus
<i>Loramycetes</i>	<i>juncicola</i>	PRJNA372853	SRR5487643	Virus
<i>Magnaporthe</i>	<i>grisea</i>	PRJNA269089	SRR1695911	Virus
<i>Morchella</i>	<i>importuna</i>	PRJNA372858	SRR5487664	Virus
<i>Neurospora</i>	<i>discreta</i>	PRJNA257829	SRR1539773	Virus
<i>Ophiocordyceps</i>	<i>sinensis</i>	PRJNA292632	SRR2533613	Virus
<i>Penicillium</i>	<i>brasilianum</i>	PRJEB7514	ERR677271	Virus
<i>Penicillium</i>	<i>digitatum</i>	PRJNA254400	SRR1557148	Virus
<i>Penicillium</i>	<i>digitatum</i>	PRJNA352307	SRR4851221	Virus
<i>Penicillium</i>	<i>raistrickii</i>	PRJNA250734	SRR1801283	Virus
<i>Periconia</i>	<i>macrospinosa</i>	PRJNA262386	SRR5153210	Virus
<i>Phyllosticta</i>	<i>citriasiatica</i>	PRJNA250394	SRR3314103	Virus
<i>Pseudogymnoascus</i>	<i>destructans</i>	PRJNA66121	SRR254216	Virus
<i>Rutstroemia</i>	<i>firma</i>	PRJNA372878	SRR5689333	Virus
<i>Sclerotinia</i>	<i>homoeocarpa</i>	PRJNA167556	SRR515157	Virus
<i>Setosphaeria</i>	<i>turcica</i>	PRJNA250530	SRR1587420	Virus
<i>Thelebolus</i>	<i>microsporus</i>	PRJNA372886	SRR5487623	Virus
<i>Tolyposcladium</i>	<i>ophioglossoides</i>	PRJNA292830	SRR2179765	Virus
<i>Trichoderma</i>	<i>asperellum</i>	PRJNA261111	SRR1575447	Virus
<i>Trichoderma</i>	<i>citrinoviride</i>	PRJNA304029	SRR2961293	Virus
<i>Trichoderma</i>	<i>harzianum</i>	PRJNA216008	SRR976276;SRR976277;SRR976278	Virus
<i>Verticillium</i>	<i>longisporum</i>	PRJNA308558	SRR3102592	Virus
<i>Zymoseptoria</i>	<i>tritici</i>	PRJEB8798	ERR789231	Virus
<i>Zymoseptoria</i>	<i>tritici</i>	PRJNA179083	SRR612175	Virus
<i>Zymoseptoria</i>	<i>tritici</i>	PRJNA237967	SRR1167717	Virus
<i>Zymoseptoria</i>	<i>tritici</i>	PRJNA253135	SRR1427070	Virus
<i>Amorphotheca</i>	<i>resinae</i>	PRJNA373008	SRR5486941	Virus (partial)
<i>Phaeosphaeriaceae</i>	sp.	PRJNA250758	SRR1587354	Virus (partial)
<i>Thozetella</i>	spPMI491	PRJNA250757	SRR1587362	Virus (partial)
<i>Ustilagoidea</i>	<i>virens</i>	PRJNA242493	SRR1203995	Virus (partial)
<i>Botrytis</i>	<i>cinerea</i>	PRJNA336478	SRR3999585	Virus (plant)
<i>Colletotrichum</i>	<i>tofieldiae</i>	PRJNA287627	SRR2072452	Virus (plant)
<i>Fusarium</i>	<i>proliferatum</i>	PRJNA392340	SRR5814194	Virus (plant)
<i>Metarhizium</i>	<i>robertsii</i>	PRJNA286035	SRR2054763	Virus (plant)
<i>Trichoderma</i>	<i>atroviride</i>	PRJNA336083	SRR3995592	Virus (plant)
<i>Trichoderma</i>	<i>reesei</i>	PRJNA282359	SRR1997914	Virus (plant)
<i>Zymoseptoria</i>	<i>brevis</i>	PRJNA277175	SRR1822514	Virus (plant)
<i>Lophiostoma</i>	<i>macrostomum</i>	PRJNA250423	SRR1797204	Virus (insect)
<i>Lophium</i>	<i>mytilinum</i>	PRJNA250592	SRR1588056	Virus (insect)

Genus	species	BioProject	SRAFile(s)	Result
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA245759	SRR1265026	Virus (insect)
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA260080	SRR1568067	Virus (insect)
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA273639	SRR1772453	Virus (insect)
<i>Pestalotiopsis</i>	<i>fici</i>	PRJNA351926	SRR4767060	Virus (insect)
<i>Xylaria</i>	<i>sp</i>	PRJNA322570	SRR3576304	Virus (insect)
<i>Aspergillus</i>	<i>ibericus</i>	PRJNA250974	SRR3501509	(contaminate)
<i>Coccomyces</i>	<i>strobi</i>	PRJNA372823	SRR5487545	(contaminate)
<i>Colletotrichum</i>	<i>navitas</i>	PRJNA262223	SRR5166244	(contaminate)
<i>Rhizina</i>	<i>undulata</i>	PRJNA372875	SRR5495983	(contaminate)
<i>Aaosphaeria</i>	<i>arxii</i>	PRJNA250693	SRR1587956	No virus
<i>Acephala</i>	<i>macrosclerotiorum</i>	PRJNA372969	SRR5487542	No virus
<i>Acremonium</i>	<i>chrysogenum</i>	PRJNA251581	SRR1339901	No virus
<i>Acremonium</i>	<i>chrysogenum</i>	PRJNA353580	SRR5020673	No virus
<i>Alternaria</i>	<i>brassicicola</i>	PRJNA156709	SRR448379	No virus
<i>Alternaria</i>	<i>brassicicola</i>	PRJNA254725	SRR1509186	No virus
<i>Amauroascus</i>	<i>mutatus</i>	PRJNA284880	SRR2045736	No virus
<i>Amauroascus</i>	<i>niger</i>	PRJNA284880	SRR2045737	No virus
<i>Ammiculicola</i>	<i>lignicola</i>	PRJNA250426	SRR1587312	No virus
<i>Amorphotheca</i>	<i>resinae</i>	PRJNA250956	SRR3314111	No virus
<i>Ampelomyces</i>	<i>quisqualis</i>	PRJNA372922	SRR5487889	No virus
<i>Anthostoma</i>	<i>avocetta</i>	PRJNA251040	SRR1796774	No virus
<i>Anthracoobia</i>	<i>macrocystis</i>	PRJNA372814	SRR5579285	No virus
<i>Arthrinium</i>	<i>malaysianum</i>	PRJNA186407	SRR746797	No virus
<i>Arthrotrichum</i>	<i>oligospora</i>	PRJNA196394	SRR824257	No virus
<i>Arthroderma</i>	<i>benhamiae</i>	PRJNA296950	SRR2560478	No virus
<i>Ascochyta</i>	<i>rabiei</i>	PRJNA288273	SRR2086542	No virus
<i>Ascodesmis</i>	<i>nigricans</i>	PRJNA259029	SRR4063760	No virus
<i>Aspergillus</i>	<i>aculeatinus</i>	PRJNA250910	SRR1799574	No virus
<i>Aspergillus</i>	<i>amstelodami</i>	PRJNA322569	SRR3576260	No virus
<i>Aspergillus</i>	<i>brasiliensis</i>	PRJNA358286	SRR5126180	No virus
<i>Aspergillus</i>	<i>brunneoviolaceus</i>	PRJNA250914	SRR1799575	No virus
<i>Aspergillus</i>	<i>calidoustus</i>	PRJEB7718	ERR677100	No virus
<i>Aspergillus</i>	<i>campestris</i>	PRJNA250800	SRR1587412	No virus
<i>Aspergillus</i>	<i>carbonarius</i>	PRJNA286841	SRR2062185	No virus
<i>Aspergillus</i>	<i>costaricaensis</i>	PRJNA250916	SRR3501512	No virus
<i>Aspergillus</i>	<i>cristatus</i>	PRJNA274642	SRR1790496	No virus
<i>Aspergillus</i>	<i>eucalypticola</i>	PRJNA250923	SRR1799576	No virus
<i>Aspergillus</i>	<i>fijiensis</i>	PRJNA250922	SRR3501506	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA144055	SRR283857	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA164607	SRR495774	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA173082	SRR544871	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA178638	SRR610538	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA189328	SRR726239	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA232295	SRR1066785	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA237320	SRR1169892	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA278877	SRR1929577	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA281697	SRR1984878	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA294865	SRR2277030	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA298505	SRR2632952	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA299060	SRR2682572	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA300619	SRR2873917	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA338480	SRR4031439	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA348383	SRR5061908	No virus
<i>Aspergillus</i>	<i>flavus</i>	PRJNA378205	SRR5314223	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJDB3180	DRR057054	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJDB4747	DRR059461	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJDB5273	DRR075868	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJEB1583	ERR236939	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJEB1586	ERR236997	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJEB3185	ERR161599	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA144647	SRR309220	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA238506	SRR1171328	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA240324	SRR1184687	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA240563	SRR1548332	No virus

Genus	species	BioProject	SRAFile(s)	Result
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA240892	SRR1186406	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA241401	SRR1197141	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA261827	SRR1583955	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA276127	SRR1813907	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA287921	SRR2076093	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA305564	SRR2984441	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA308356	SRR3095794	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA308518	SRR3098040	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA312558	SRR3233134	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA323497	SRR3591503	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA327227	SRR3734545	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA336568	SRR4000584	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA382191	SRR5441597	No virus
<i>Aspergillus</i>	<i>fumigatus</i>	PRJNA52787	SRR068954	No virus
<i>Aspergillus</i>	<i>glaucus</i>	PRJNA291955	SRR2163054	No virus
<i>Aspergillus</i>	<i>indologenus</i>	PRJNA250978	SRR1801410	No virus
<i>Aspergillus</i>	<i>japonicus</i>	PRJNA250976	SRR4063263	No virus
<i>Aspergillus</i>	<i>lacticoffeatus</i>	PRJNA250971	SRR1799579	No virus
<i>Aspergillus</i>	<i>luchuensis</i>	PRJNA356975	SRR5097374	No virus
<i>Aspergillus</i>	<i>luchuensis</i>	PRJNA358288	SRR5126146	No virus
<i>Aspergillus</i>	<i>montevidensis</i>	PRJNA322569	SRR3576259	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJDB3191	DRR059453	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJDB4821	DRR060260	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA188657	SRR676520	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA268669	SRR1664216	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA276127	SRR1816788	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA293111	SRR2170285	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA293709	SRR2180251	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA294437	SRR2225887	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA311945	SRR3212502	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA337983	SRR4009440	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA345604	SRR4368910	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA355963	SRR5069240	No virus
<i>Aspergillus</i>	<i>nidulans</i>	PRJNA373914	SRR5241853	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJDB4747	DRR059467	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA205471	SRR867792	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA250529	SRR1799278	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA255841	SRR1524306	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA299483	SRR2771682	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA318352	SRR3374423	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA320850	SRR3479369	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA327096	SRR3726215	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA328955	SRR3921618	No virus
<i>Aspergillus</i>	<i>niger</i>	PRJNA350271	SRR4446735	No virus
<i>Aspergillus</i>	<i>novofumigatus</i>	PRJNA250796	SRR1587370	No virus
<i>Aspergillus</i>	<i>ochraceoroseus</i>	PRJNA250798	SRR1587379	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJDB1880	DRR003571	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJDB2139	DRR003569	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJDB2320	DRR003570	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJDB2646	DRR003573	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJDB2791	DRR003572	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJDB4747	DRR059464	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJNA121091	SRR063693	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJNA178638	SRR610542	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJNA248324	SRR1425374	No virus
<i>Aspergillus</i>	<i>oryzae</i>	PRJNA305561	SRR2984402	No virus
<i>Aspergillus</i>	<i>piperis</i>	PRJNA250968	SRR4063260	No virus
<i>Aspergillus</i>	<i>saccharolyticus</i>	PRJNA250985	SRR1799580	No virus
<i>Aspergillus</i>	<i>scleroticarbonarius</i>	PRJNA250912	SRR1801401	No virus
<i>Aspergillus</i>	<i>sclerotioniger</i>	PRJNA250913	SRR1799563	No virus
<i>Aspergillus</i>	<i>steynii</i>	PRJNA250799	SRR1587405	No virus
<i>Aspergillus</i>	<i>sydowii</i>	PRJNA207689	SRR486614	No virus
<i>Aspergillus</i>	<i>sydowii</i>	PRJNA242454	SRR1201684	No virus
<i>Aspergillus</i>	<i>sydowii</i>	PRJNA304078	SRR2960314	No virus

Genus	species	BioProject	SRAFile(s)	Result
<i>Aspergillus</i>	<i>sydowii</i>	PRJNA353756	SRR5576788	No virus
<i>Aspergillus</i>	<i>sydowii</i>	PRJNA358293	SRR5126119	No virus
<i>Aspergillus</i>	<i>terreus</i>	PRJNA296012	SRR2409422	No virus
<i>Aspergillus</i>	<i>terreus</i>	PRJNA360953	SRR5167787	No virus
<i>Aspergillus</i>	<i>tubingensis</i>	PRJNA358285	SRR5126196	No virus
<i>Aspergillus</i>	<i>uvarum</i>	PRJNA250981	SRR1799582	No virus
<i>Aspergillus</i>	<i>vadensis</i>	PRJNA250977	SRR1799583	No virus
<i>Aspergillus</i>	<i>versicolor</i>	PRJNA358291	SRR5126117	No virus
<i>Aspergillus</i>	<i>violaceofuscus</i>	PRJNA250975	SRR3501508	No virus
<i>Aspergillus</i>	<i>wentii</i>	PRJNA358290	SRR5126179	No virus
<i>Aureobasidium</i>	<i>melanogenum</i>	PRJNA309099	SRR3109571	No virus
<i>Aureobasidium</i>	<i>pullulans</i>	PRJNA251264	SRR5145577	No virus
<i>Aureobasidium</i>	<i>pullulans</i>	PRJNA301913	SRR2915685	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA174866	SRR567506	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA193558	SRR786686	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA232447	SRR1057281	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA253184	SRR1427291	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA226222	SRR1023677	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA257254	SRR1534388	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA273183	SRR1767720	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA291528	SRR2136033	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA383131	SRR5453675	No virus
<i>Beauveria</i>	<i>bassiana</i>	PRJNA173231	SRR545683	No virus
<i>Bimuria</i>	<i>novae-zelandiae</i>	PRJNA250572	SRR1588024	No virus
<i>Bipolaris</i>	<i>sorokiniana</i>	PRJNA53923	SRR3927048	No virus
<i>Botrytis</i>	<i>cinerea</i>	PRJNA293283	SRR2174493	No virus
<i>Botrytis</i>	<i>cinerea</i>	PRJNA351919	SRR5040505	No virus
<i>Buergenerula</i>	<i>spartinae</i>	PRJNA269089	SRR1695860	No virus
<i>Bulgaria</i>	<i>inquinans</i>	PRJNA372815	SRR5488107	No virus
<i>Bussabanomyces</i>	<i>longisporus</i>	PRJNA269089	SRR1695861	No virus
<i>Byssosnygena</i>	<i>ceratinophila</i>	PRJNA284880	SRR2045738	No virus
<i>Byssothecium</i>	<i>circinans</i>	PRJNA250584	SRR3439938	No virus
<i>Cadophora</i>	<i>sp.</i>	PRJNA250852	SRR1587637	No virus
<i>Caloscypha</i>	<i>fulgens</i>	PRJNA372816	SRR5491274	No virus
<i>Cenococcum</i>	<i>geophilum</i>	PRJNA196023	SRR398303	No virus
<i>Cenococcum</i>	<i>geophilum</i>	PRJNA327385	SRR3734050	No virus
<i>Cercospora</i>	<i>zeina</i>	PRJNA355557	SRR5063175	No virus
<i>Chalara</i>	<i>longipes</i>	PRJNA251037	SRR1796783	No virus
<i>Choiromyces</i>	<i>venosus</i>	PRJNA250524	SRR1587749	No virus
<i>Chrysosporium</i>	<i>queenslandicum</i>	PRJNA284880	SRR2045740	No virus
<i>Cladophialophora</i>	<i>immunda</i>	PRJNA325796	SRR2571134	No virus
<i>Clathrospora</i>	<i>elynae</i>	PRJNA259052	SRR4054983	No virus
<i>Clonostachys</i>	<i>rosea</i>	PRJNA258670	SRR4063417	No virus
<i>Clonostachys</i>	<i>rosea</i>	PRJNA272346	SRR1751511	No virus
<i>Coccidioides</i>	<i>posadasii</i>	PRJNA338286	SRR4015208	No virus
<i>Coccolini</i>	<i>bartschii</i>	PRJNA259025	SRR4052162	No virus
<i>Colletotrichum</i>	<i>falcatum</i>	PRJNA262215	SRR5166051	No virus
<i>Colletotrichum</i>	<i>fioriniae</i>	PRJNA372827	SRR5506722	No virus
<i>Colletotrichum</i>	<i>gloeosporioides</i>	PRJNA178310	SRR609271	No virus
<i>Colletotrichum</i>	<i>gloeosporioides</i>	PRJNA238599	SRR1171641	No virus
<i>Colletotrichum</i>	<i>gloeosporioides</i>	PRJNA300348	SRR2844617	No virus
<i>Colletotrichum</i>	<i>gloeosporioides</i>	PRJNA391239	SRR5723443	No virus
<i>Colletotrichum</i>	<i>godetiae</i>	PRJNA372825	SRR5506725	No virus
<i>Colletotrichum</i>	<i>incanum</i>	PRJNA287627	SRR2072467	No virus
<i>Colletotrichum</i>	<i>lupini</i>	PRJNA372826	SRR5506723	No virus
<i>Colletotrichum</i>	<i>somersetensis</i>	PRJNA262443	SRR5166063	No virus
<i>Colletotrichum</i>	<i>sublineola</i>	PRJNA262379	SRR5145573	No virus
<i>Coniochaeta</i>	<i>lignaria</i>	PRJNA256711	SRR2131145	No virus
<i>Coniochaeta</i>	<i>lignaria</i>	PRJNA256712	SRR1588071	No virus
<i>Coniochaeta</i>	<i>spPMI546</i>	PRJNA250756	SRR3453022	No virus
<i>Cordyceps</i>	<i>cicadae</i>	PRJNA324675	SRR3655293	No virus
<i>Cordyceps</i>	<i>cicadae</i>	PRJNA376476	SRR5279602	No virus
<i>Cordyceps</i>	<i>militaris</i>	PRJNA170029	SRR518537	No virus
<i>Corollospora</i>	<i>maritima</i>	PRJNA250801	SRR3490952	No virus

Genus	species	BioProject	SRAFile(s)	Result
<i>Corollospora</i>	<i>maritima</i>	PRJNA274818	SRR1793202	No virus
<i>Corynespora</i>	<i>cassicola</i>	PRJNA250634	SRR4061751	No virus
<i>Cryphonectria</i>	<i>parasitica</i>	PRJNA258639	SRR1557137	No virus
<i>Cryphonectria</i>	<i>parasitica</i>	PRJNA369604	SRR5235468	No virus
<i>Daldinia</i>	<i>eschscholtzii</i>	PRJNA250743	SRR1798141	No virus
<i>Decorospora</i>	<i>gaudefroyi</i>	PRJNA372915	SRR5487497	No virus
<i>Diatrype</i>	<i>disciformis</i>	PRJNA251269	SRR1797162	No virus
<i>Didymella</i>	<i>exigua</i>	PRJNA207857	SRR488350	No virus
<i>Didymocrea</i>	<i>sadasivanii</i>	PRJNA372835	SRR5494005	No virus
<i>Didymosphaeria</i>	<i>enalia</i>	PRJNA256706	SRR1587972	No virus
<i>Dissoconium</i>	<i>aciculare</i>	PRJNA250413	SRR1797190	No virus
<i>Dothidotthia</i>	<i>symphoricarpi</i>	PRJNA250424	SRR1587320	No virus
<i>Drechmeria</i>	<i>coniospora</i>	PRJNA269584	SRR1930119	No virus
<i>Epichloe</i>	<i>festucae</i>	PRJNA299335	SRR3102671	No virus
<i>Epichloe</i>	<i>festucae</i>	PRJNA308860	SRR3105525	No virus
<i>Epichloe</i>	<i>hybrida</i>	PRJNA299335	SRR3102659	No virus
<i>Epichloe</i>	<i>typhina</i>	PRJNA299335	SRR3102668	No virus
<i>Epichloe</i>	<i>typhina</i>	PRJNA308860	SRR3105526	No virus
<i>Eremomyces</i>	<i>bilateralis</i>	PRJNA250468	SRR1587346	No virus
<i>Escovopsis</i>	<i>weberi</i>	PRJNA253870	SRR1639281	No virus
<i>Exophiala</i>	<i>mesophila</i>	PRJNA325797	SRR2570632	No virus
<i>Exophiala</i>	<i>oligosperma</i>	PRJNA325798	SRR2570634	No virus
<i>Exophiala</i>	<i>sideris</i>	PRJNA325799	SRR2570635	No virus
<i>Exophiala</i>	<i>spinifera</i>	PRJNA325800	SRR2571135	No virus
<i>Exophiala</i>	<i>xenobiotica</i>	PRJNA325801	SRR2570633	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA197332	SRR828733	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA221153	SRR999638	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA227143	SRR1027330	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA239711	SRR1179895	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA240256	SRR1185285	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA254543	SRR1508760	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA262571	SRR1592424	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA262846	SRR1596060	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA263651	SRR4445676	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA273653	SRR1772867	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA289285	SRR2518132	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA292945	SRR3168567	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA293109	SRR2170115	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA293594	SRR2182495	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA298346	SRR2603147	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA304218	SRR3055829	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA314297	SRR3203806	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA316100	SRR3286072	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA326901	SRR3721336	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA345628	SRR4372406	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA347833	SRR4415737	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA376074	SRR5282269	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA376640	SRR5282546	No virus
<i>Fusarium</i>	<i>graminearum</i>	PRJNA390225	SRR5678737	No virus
<i>Fusarium</i>	<i>oxysporum</i>	PRJNA294248	SRR3474099	No virus
<i>Fusarium</i>	<i>oxysporum</i>	PRJNA303949	SRR2960134	No virus
<i>Fusarium</i>	<i>oxysporum</i>	PRJNA322439	SRR3727692	No virus
<i>Fusarium</i>	<i>oxysporum</i>	PRJNA323849	SRR3609438	No virus
<i>Fusarium</i>	<i>oxysporum</i>	PRJNA48015	SRR190803	No virus
<i>Fusarium</i>	<i>proliferatum</i>	PRJNA311604	SRR3161849	No virus
<i>Fusarium</i>	<i>pseudograminearum</i>	PRJNA326033	SRR3695294	No virus
<i>Fusarium</i>	<i>solani</i>	PRJNA323849	SRR3609447	No virus
<i>Fusarium</i>	<i>verticillioides</i>	PRJNA262571	SRR1592418	No virus
<i>Fusarium</i>	<i>verticillioides</i>	PRJNA275769	SRR1810215	No virus
<i>Fusarium</i>	<i>verticillioides</i>	PRJNA320757	SRR3479291	No virus
<i>Fusarium</i>	<i>virguliforme</i>	PRJNA248018	SRR1382101	No virus
<i>Fusarium</i>	<i>virguliforme</i>	PRJNA289542	SRR3193213	No virus
<i>Gaeumannomyces</i>	<i>avenae</i>	PRJNA269089	SRR1695862	No virus
<i>Gaeumannomyces</i>	<i>graminis</i>	PRJNA269089	SRR1695910	No virus

Genus	species	BioProject	SRAFile(s)	Result
<i>Gaeumannomyces</i>	<i>radicicola</i>	PRJNA269089	SRR1695864	No virus
<i>Gaeumannomyces</i>	<i>tritici</i>	PRJNA52815	SRR081550	No virus
<i>Grossmannia</i>	<i>clavigera</i>	PRJNA39837	SRR037886	No virus
<i>Gymnascella</i>	<i>aurantiaca</i>	PRJNA196037	SRR497474	No virus
<i>Histoplasma</i>	<i>capsulatum</i>	PRJNA214093	SRR949916	No virus
<i>Histoplasma</i>	<i>capsulatum</i>	PRJNA283425	SRR2015209	No virus
<i>Hortaea</i>	<i>acidophila</i>	PRJNA372841	SRR5487552	No virus
<i>Hypoxylon</i>	<i>sp.</i>	PRJNA250742	SRR1801292	No virus
<i>Hypoxylon</i>	<i>spEC38</i>	PRJNA250737	SRR1798132	No virus
<i>Isaria</i>	<i>catenianulata</i>	PRJNA317747	SRR3501158	No virus
<i>Kalaharituber</i>	<i>pfeilli</i>	PRJNA372846	SRR5506696	No virus
<i>Karstenula</i>	<i>rhodostoma</i>	PRJNA250576	SRR1588053	No virus
<i>Khuskia</i>	<i>oryzae</i>	PRJNA372847	SRR5489234	No virus
<i>Knoxdaviesia</i>	<i>proteae</i>	PRJNA275563	SRR2962986	No virus
<i>Lasallia</i>	<i>pustulata</i>	PRJEB11664	ERR1110285	No virus
<i>Lasiodiplodia</i>	<i>theobromae</i>	PRJNA305901	SRR2994047	No virus
<i>Lasiodiplodia</i>	<i>theobromae</i>	PRJNA387508	SRR5585649	No virus
<i>Leptosphaeria</i>	<i>biglobosa</i>	PRJNA230885	SRR1151408	No virus
<i>Leptosphaeria</i>	<i>maculans</i>	PRJNA230885	SRR1151407	No virus
<i>Leptosphaeria</i>	<i>maculans</i>	PRJNA322501	SRR3571668	No virus
<i>Leptosphaeria</i>	<i>maculans</i>	PRJNA328045	SRR3750692	No virus
<i>Lindgomyces</i>	<i>ingoldianus</i>	PRJNA372850	SRR5487617	No virus
<i>Lineolata</i>	<i>rhizophorae</i>	PRJNA372851	SRR5489199	No virus
<i>Lollipopaia</i>	<i>minuta</i>	PRJNA372916	SRR5487522	No virus
<i>Lophodemium</i>	<i>nitens</i>	PRJNA262384	SRR5153207	No virus
<i>Loramycetes</i>	<i>macrosporus</i>	PRJNA372852	SRR5488109	No virus
<i>Macgarvieomyces</i>	<i>juncicola</i>	PRJNA269089	SRR1695865	No virus
<i>Macrophomina</i>	<i>phaseolina</i>	PRJNA326815	SRR5282579	No virus
<i>Macroventuria</i>	<i>anomochaeta</i>	PRJNA250579	SRR1587965	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJDB2669	DRR001967	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJDB2912	DRR021448	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA143617	SRR306516	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA187572	SRR653484	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA209870	SRR923948	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA223667	SRR1015598	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA259589	SRR1561422	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA288294	SRR2079616	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA301902	SRR2932784	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA312478	SRR3176886	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA322180	SRR3545708	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA52817	SRR081552	No virus
<i>Magnaporthe</i>	<i>oryzae</i>	PRJNA187047	SRR651979	No virus
<i>Magnaporthiopsis</i>	<i>incrustans</i>	PRJNA269089	SRR1692485	No virus
<i>Magnaporthiopsis</i>	<i>panicorum</i>	PRJNA269089	SRR1695868	No virus
<i>Magnaporthiopsis</i>	<i>poae</i>	PRJNA48987	SRR051974	No virus
<i>Magnaporthiopsis</i>	<i>rhizophila</i>	PRJNA269089	SRR1695856	No virus
<i>Massarina</i>	<i>ebumea</i>	PRJNA250462	SRR1587358	No virus
<i>Microascus</i>	<i>trigonosporus</i>	PRJNA259026	SRR4052567	No virus
<i>Microdochium</i>	<i>bolleyi</i>	PRJNA372857	SRR5487662	No virus
<i>Microthyrium</i>	<i>microscopicum</i>	PRJNA259047	SRR1587365	No virus
<i>Monascus</i>	<i>purpureus</i>	PRJNA204098	SRR492869	No virus
<i>Monilinia</i>	<i>fructicola</i>	PRJNA255662	SRR1521463	No virus
<i>Morchella</i>	<i>conica</i>	PRJNA250804	SRR1799172	No virus
<i>Myceliophthora</i>	<i>thermophila</i>	PRJNA137175	SRR101406	No virus
<i>Myrothecium</i>	<i>inundatum</i>	PRJNA259049	SRR4096435	No virus
<i>Nakataea</i>	<i>oryzae</i>	PRJNA269089	SRR1695857	No virus
<i>Neofusicoccum</i>	<i>parvum</i>	PRJNA336064	SRR3992649	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA152849	SRR400659	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA153791	SRR447484	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA182734	SRR627947	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA188720	SRR679736	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA191575	SRR765751	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA217182	SRR957218	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA227035	SRR1025973	No virus

Genus	species	BioProject	SRAFile(s)	Result
<i>Neurospora</i>	<i>crassa</i>	PRJNA230677	SRR1043589	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA230720	SRR1043827	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA232306	SRR1055985	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA248256	SRR4175080	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA250475	SRR1570816	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA254459	SRR1538716	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA256749	SRR5133932	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA257279	SRR1534767	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA258518	SRR1578070	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA260043	SRR1562554	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA262776	SRR1593998	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA282545	SRR2036201	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA282963	SRR2007029	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA290152	SRR2105977	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA303140	SRR2952641	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA312794	SRR3181960	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA314110	SRR5204538	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA314494	SRR3207376	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA324378	SRR3623688	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA339581	SRR4045330	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA347916	SRR4431501	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA349561	SRR5065838	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA352692	SRR5000482	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA357766	SRR1588293	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA378486	SRR5320490	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA388208	SRR5621551	No virus
<i>Neurospora</i>	<i>crassa</i>	PRJNA72789	SRR341283	No virus
<i>Neurospora</i>	<i>tetrasperma</i>	PRJNA257828	SRR1539785	No virus
<i>Niesslia</i>	<i>exilis</i>	PRJNA259055	SRR4061764	No virus
<i>Oidiodendron</i>	<i>maius</i>	PRJNA269503	SRR1695518	No virus
<i>Omnidemptus</i>	<i>affinis</i>	PRJNA269089	SRR1695873	No virus
<i>Ophiobolus</i>	<i>disseminans</i>	PRJNA250589	SRR1588061	No virus
<i>Ophioceras</i>	<i>commune</i>	PRJNA269089	SRR1695874	No virus
<i>Ophioceras</i>	<i>dolichostomum</i>	PRJNA269089	SRR1695875	No virus
<i>Ophioceras</i>	<i>leptosporum</i>	PRJNA269089	SRR1695877	No virus
<i>Ophiocordyceps</i>	<i>sinensis</i>	PRJNA271505	SRR1742983	No virus
<i>Ophiocordyceps</i>	<i>sinensis</i>	PRJNA308334	SRR3095647	No virus
<i>Ophiocordyceps</i>	<i>sinensis</i>	PRJNA325365	SRR3658816	No virus
<i>Ophiocordyceps</i>	<i>sinensis</i>	PRJNA376530	SRR5282569	No virus
<i>Ophiocordyceps</i>	<i>sinensis</i>	PRJNA382001	SRR5428527	No virus
<i>Ophiocordyceps</i>	<i>unilateralis</i>	PRJNA281955	SRR1986857	No virus
<i>Ophiostoma</i>	<i>novo ulmi</i>	PRJNA260920	SRR2140676	No virus
<i>Paecilomyces</i>	<i>hepiali</i>	PRJNA203663	SRR900012	No virus
<i>Paracoccidioides</i>	<i>lutzii</i>	PRJNA297276	SRR2547555	No virus
<i>Paraphaeosphaeria</i>	<i>sporulosa</i>	PRJNA259139	SRR4063290	No virus
<i>Passalora</i>	<i>fulva</i>	PRJNA86753	SRR1171046	No virus
<i>Penicillium</i>	<i>bilaiiae</i>	PRJNA373026	SRR5487470	No virus
<i>Penicillium</i>	<i>bilaiiae</i>	PRJNA403282	SRR6031138	No virus
<i>Penicillium</i>	<i>brevicompactum</i>	PRJNA250741	SRR1801281	No virus
<i>Penicillium</i>	<i>brevicompactum</i>	PRJNA256758	SRR3502875	No virus
<i>Penicillium</i>	<i>canescens</i>	PRJNA373028	SRR5487465	No virus
<i>Penicillium</i>	<i>capsulatum</i>	PRJNA338896	SRR4031065	No virus
<i>Penicillium</i>	<i>capsulatum</i>	PRJNA339902	SRR4051963	No virus
<i>Penicillium</i>	<i>chrysogenum</i>	PRJNA270038	SRR1705825	No virus
<i>Penicillium</i>	<i>chrysogenum</i>	PRJNA287440	SRR2070795	No virus
<i>Penicillium</i>	<i>chrysogenum</i>	PRJNA288074	SRR2079293	No virus
<i>Penicillium</i>	<i>chrysogenum</i>	PRJNA320850	SRR3479402	No virus
<i>Penicillium</i>	<i>chrysogenum</i>	PRJNA353579	SRR5020747	No virus
<i>Penicillium</i>	<i>expansum</i>	PRJNA222879	SRR1508205	No virus
<i>Penicillium</i>	<i>expansum</i>	PRJNA314168	SRR3201551	No virus
<i>Penicillium</i>	<i>expansum</i>	PRJNA373025	SRR5487466	No virus
<i>Penicillium</i>	<i>expansum</i>	PRJNA403270	SRR6031084	No virus
<i>Penicillium</i>	<i>fellutanum</i>	PRJNA373029	SRR5487468	No virus
<i>Penicillium</i>	<i>glabrum</i>	PRJNA373027	SRR5487463	No virus

Genus	species	BioProject	SRAFile(s)	Result
<i>Penicillium</i>	<i>janthinellum</i>	PRJNA250735	SRR1801286	No virus
<i>Penicillium</i>	<i>janthinellum</i>	PRJNA259826	SRR1639172	No virus
<i>Penicillium</i>	<i>lanosocoeruleum</i>	PRJNA250525	SRR1797397	No virus
<i>Penicillium</i>	<i>oxalicum</i>	PRJNA223338	SRR1022686	No virus
<i>Penicillium</i>	<i>oxalicum</i>	PRJNA285095	SRR2042781	No virus
<i>Penicillium</i>	<i>oxalicum</i>	PRJNA294261	SRR2217919	No virus
<i>Penicillium</i>	<i>oxalicum</i>	PRJNA299612	SRR2758969	No virus
<i>Penicillium</i>	<i>oxalicum</i>	PRJNA316645	SRR3308142	No virus
<i>Penicillium</i>	<i>oxalicum</i>	PRJNA317762	SRR3350644	No virus
<i>Penicillium</i>	<i>oxalicum</i>	PRJNA331129	SRR3948934	No virus
<i>Pestalotiopsis</i>	<i>fici</i>	PRJNA257381	SRR1535667	No virus
<i>Pezoloma</i>	<i>ericae</i>	PRJNA259018	SRR4051865	No virus
<i>Phaeoacremonium</i>	<i>minimum</i>	PRJNA279250	SRR1927072	No virus
<i>Phaeoconiella</i>	<i>chlamydospora</i>	PRJNA279250	SRR1927076	No virus
<i>Phialocephala</i>	<i>scopiformis</i>	PRJNA265599	SRR4063275	No virus
<i>Phialocephala</i>	<i>subalpina</i>	PRJEB12610	ERR1233862	No virus
<i>Phialophora</i>	<i>americana</i>	PRJNA325795	SRR2570630	No virus
<i>Phomopsis</i>	<i>sp.</i>	PRJNA376069	SRR5279398	No virus
<i>Plectania</i>	<i>melastoma</i>	PRJNA372869	SRR5487549	No virus
<i>Plectosphaerella</i>	<i>cucumerina</i>	PRJNA259141	SRR4063267	No virus
<i>Plenodomus</i>	<i>tracheiphilus</i>	PRJNA259093	SRR3452797	No virus
<i>Podospora</i>	<i>anserina</i>	PRJNA313943	SRR3197700	No virus
<i>Polyplosphaeria</i>	<i>fusca</i>	PRJNA259053	SRR4061903	No virus
<i>Pseudocercospora</i>	<i>fijiensis</i>	PRJNA323552	SRR3593877	No virus
<i>Pseudogymnoascus</i>	<i>destructans</i>	PRJNA246154	SRR1270417	No virus
<i>Pseudogymnoascus</i>	<i>destructans</i>	PRJNA246155	SRR1270712	No virus
<i>Pseudohalonectria</i>	<i>lignicola</i>	PRJNA269089	SRR1695878	No virus
<i>Pseudomassariella</i>	<i>vexata</i>	PRJNA372873	SRR5494126	No virus
<i>Pseudophialophora</i>	<i>eragrostis</i>	PRJNA269089	SRR1695880	No virus
<i>Pseudophialophora</i>	<i>panicorum</i>	PRJNA269089	SRR1695881	No virus
<i>Pseudophialophora</i>	<i>schizachyrii</i>	PRJNA269089	SRR1695882	No virus
<i>Pseudosarcosoma</i>	<i>latahense</i>	PRJNA372874	SRR5487621	No virus
<i>Pseudovirgaria</i>	<i>hyperparasitica</i>	PRJNA262391	SRR5165209	No virus
<i>Pyrenochaeta</i>	<i>lycopersica</i>	PRJNA202292	SRR851823	No virus
<i>Pyrenochaeta</i>	<i>sp.</i>	PRJNA259142	SRR4063377	No virus
<i>Pyrenophora</i>	<i>teres</i>	PRJEB18107	ERR1743401	No virus
<i>Pyronema</i>	<i>omphalodes</i>	PRJNA177769	SRR593254	No virus
<i>Pyronema</i>	<i>omphalodes</i>	PRJNA260620	SRR1571977	No virus
<i>Reticulascus</i>	<i>tulasneorum</i>	PRJNA310383	SRR3138063	No virus
<i>Rhizodiscina</i>	<i>lignyota</i>	PRJNA250698	SRR1588057	No virus
<i>Sarcoscypha</i>	<i>coccinea</i>	PRJNA372879	SRR5487620	No virus
<i>Sarocladium</i>	<i>strictum</i>	PRJNA259144	SRR4063304	No virus
<i>Sclerotinia</i>	<i>sclerotiorum</i>	PRJNA261444	SRR1582469	No virus
<i>Sclerotinia</i>	<i>sclerotiorum</i>	PRJNA327437	SRR3734873	No virus
<i>Sclerotinia</i>	<i>sclerotiorum</i>	PRJNA362796	SRR5199360	No virus
<i>Scopulariopsis</i>	<i>brevicaulis</i>	PRJNA260949	SRR1578116	No virus
<i>Shiraia</i>	<i>bambusicola</i>	PRJNA292353	SRR2153022	No virus
<i>Shiraia</i>	<i>bambusicola</i>	PRJNA323638	SRR3602718	No virus
<i>Slopeiomyces</i>	<i>cylindrosporus</i>	PRJNA269089	SRR1695908	No virus
<i>Sordaria</i>	<i>macrospora</i>	PRJNA148727	SRR364309	No virus
<i>Sordaria</i>	<i>macrospora</i>	PRJNA213728	SRR944971	No virus
<i>Sphaerulina</i>	<i>populicola</i>	PRJNA277053	SRR1797339	No virus
<i>Sporomia</i>	<i>finetaria</i>	PRJNA250425	SRR1797186	No virus
<i>Stagonospora</i>	<i>sp.</i>	PRJNA259143	SRR4063378	No virus
<i>Stanjemonium</i>	<i>grisellum</i>	PRJNA259148	SRR4063396	No virus
<i>Talaromyces</i>	<i>aculeatus</i>	PRJNA250738	SRR1801282	No virus
<i>Talaromyces</i>	<i>mameffei</i>	PRJNA207279	SRR922412	No virus
<i>Talaromyces</i>	<i>mameffei</i>	PRJNA212740	SRR941608	No virus
<i>Talaromyces</i>	<i>mameffei</i>	PRJNA251718	SRR1514345	No virus
<i>Talaromyces</i>	<i>purpureogenus</i>	PRJNA276974	SRR1824012	No virus
<i>Teratosphaeria</i>	<i>nubilosa</i>	PRJNA259048	SRR4061811	No virus
<i>Terfezia</i>	<i>boudieri</i>	PRJNA258988	SRR4051876	No virus
<i>Thermomyces</i>	<i>lanuginosus</i>	PRJNA213840	SRR946591	No virus
<i>Thermomyces</i>	<i>lanuginosus</i>	PRJNA286684	SRR2059702	No virus

Genus	species	BioProject	SRAFile(s)	Result
<i>Thermothelomyces</i>	<i>heterothallica</i>	PRJNA373016	SRR5487449	No virus
<i>Thielavia</i>	<i>antarctica</i>	PRJNA250591	SRR3439670	No virus
<i>Thielavia</i>	<i>appendiculata</i>	PRJNA250578	SRR3439582	No virus
<i>Thielavia</i>	<i>arenaria</i>	PRJNA250583	SRR3439600	No virus
<i>Thielavia</i>	<i>hyrcaniae</i>	PRJNA250619	SRR3439602	No virus
<i>Thielavia</i>	<i>subthermophila</i>	PRJNA385082	SRR5527128	No virus
<i>Thielavia</i>	<i>terrestris</i>	PRJNA137175	SRR101413	No virus
<i>Togninia</i>	<i>minima</i>	PRJNA270932	SRR1729158	No virus
<i>Tolypocladium</i>	<i>inflatum</i>	PRJNA385765	SRR5515128	No virus
<i>Tothia</i>	<i>fuscella</i>	PRJNA250692	SRR3439940	No virus
<i>Trematosphaeria</i>	<i>pertusa</i>	PRJNA250571	SRR1588067	No virus
<i>Trichodelitschia</i>	<i>bisporula</i>	PRJNA259056	SRR4061759	No virus
<i>Trichoderma</i>	<i>asperellum</i>	PRJNA207877	SRR394065	No virus
<i>Trichoderma</i>	<i>asperellum</i>	PRJNA279109	SRR1926139	No virus
<i>Trichoderma</i>	<i>atroviride</i>	PRJNA310123	SRR2230025	No virus
<i>Trichoderma</i>	<i>brevicompectum</i>	PRJNA183441	SRR630525	No virus
<i>Trichoderma</i>	<i>citrinoviride</i>	PRJNA291665	SRR2148793	No virus
<i>Trichoderma</i>	<i>harzianum</i>	PRJNA175485	SRR579379	No virus
<i>Trichoderma</i>	<i>harzianum</i>	PRJNA207867	SRR387795	No virus
<i>Trichoderma</i>	<i>harzianum</i>	PRJNA272748	SRR1765897	No virus
<i>Trichoderma</i>	<i>longibrachiatum</i>	PRJNA207876	SRR391286	No virus
<i>Trichoderma</i>	<i>longibrachiatum</i>	PRJNA242904	SRR1207063	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA213530	SRR944617	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA230357	SRR1045101	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA232512	SRR1057954	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA246360	SRR1273834	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA255678	SRR1520706	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA278361	SRR1916249	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA278646	SRR1917215	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA293671	SRR2192595	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA320850	SRR3479383	No virus
<i>Trichoderma</i>	<i>reesei</i>	PRJNA372889	SRR5506695	No virus
<i>Trichophyton</i>	<i>rubrum</i>	PRJNA276016	SRR1812769	No virus
<i>Trichophyton</i>	<i>rubrum</i>	PRJNA276019	SRR1812771	No virus
<i>Trichophyton</i>	<i>rubrum</i>	PRJNA276041	SRR1812893	No virus
<i>Trichophyton</i>	<i>rubrum</i>	PRJNA276042	SRR1812894	No virus
<i>Trichophyton</i>	<i>rubrum</i>	PRJNA88099	SRR1596318	No virus
<i>Trinosporium</i>	<i>guianense</i>	PRJNA258710	SRR4125813	No virus
<i>Tuber</i>	<i>borchii</i>	PRJNA384618	SRR5623538	No virus
<i>Tuber</i>	<i>melanosporum</i>	PRJNA214782	SRR1237938	No virus
<i>Usnea</i>	<i>florida</i>	PRJNA372893	SRR5489198	No virus
<i>Ustilagoidea</i>	<i>virens</i>	PRJNA344467	SRR4296235	No virus
<i>Valettoniellopsis</i>	<i>laxa</i>	PRJNA259051	SRR4061781	No virus
<i>Valsa</i>	<i>mali</i>	PRJNA232503	SRR1060361	No virus
<i>Venturia</i>	<i>inaequalis</i>	PRJNA292699	SRR2164202	No virus
<i>Verruconis</i>	<i>gallopava</i>	PRJNA325802	SRR2570631	No virus
<i>Verticillium</i>	<i>dahliae</i>	PRJNA169154	SRR575673	No virus
<i>Verticillium</i>	<i>dahliae</i>	PRJNA196692	SRR824366	No virus
<i>Verticillium</i>	<i>dahliae</i>	PRJNA208120	SRR1232601	No virus
<i>Verticillium</i>	<i>dahliae</i>	PRJNA288926	SRR2087158	No virus
<i>Verticillium</i>	<i>dahliae</i>	PRJNA381107	SRR5440696	No virus
<i>Verticillium</i>	<i>nonalfalfae</i>	PRJNA283258	SRR2012820	No virus
<i>Westerdykella</i>	<i>omata</i>	PRJNA250582	SRR1588066	No virus
<i>Xenopyricularia</i>	<i>zizaniicola</i>	PRJNA269089	SRR1695886	No virus
<i>Xylaria</i>	<i>striata</i>	PRJNA304370	SRR2962744	No virus
<i>Xylona</i>	<i>heveae</i>	PRJNA250632	SRR1798062	No virus
<i>Zymoseptoria</i>	<i>ardabiliae</i>	PRJNA259097	SRR4097075	No virus
<i>Zymoseptoria</i>	<i>ardabiliae</i>	PRJNA277174	SRR1822499	No virus
<i>Zymoseptoria</i>	<i>pseudotritici</i>	PRJNA277173	SRR1822493	No virus

Table S2: Name and Accession IDs of mycoviruses related to those identified in this study

Virus Name	Virus Abbrev.	NCBI ID	Used for analysis of:
Alternaria longipes dsRNA virus 1	AIV1-HN28	YP_009052469.1	Non-segmented viruses
Beauveria bassiana RNA virus 1	BbV1	YP_009154711.1	Non-segmented viruses
Beauveria bassiana RNA virus 1	BbV1-A24	AKC57301.1	Non-segmented viruses
Penicillium janczewskii Beauveria bassiana-like virus 1	PjbbV1	ALO50135.1	Non-segmented viruses
Ustilagoidea vires unassigned RNA virus HNND-1	UvV-HNND1	YP_009154709.1	Non-segmented viruses
Alternaria alternata virus 1	AaV1	YP_001976142	Altemaviridae
Aspergillus foetidus dsRNA mycovirus	AfdsV	YP_007353985	Altemaviridae
Fusarium poae altamavirus 1	FpAV1	NC_030883.1	Altemaviridae
Fusarium graminearum altamavirus 1	FgAV1	YP_009449439	Altemaviridae
Aspergillus (niger) mycovirus 341	AnMV341	ABX79997	Altemaviridae
Aspergillus fumigatus chrysovirus	AFV	FN178512, FN178513, FN178514, FN178515	Chrysoviridae
Penicillium chrysogenum virus	PcCV	AF296439, AF296440, AF296441, AF296442	Chrysoviridae
Verticillium dahliae chrysovirus 1	VdCV1	HM004067, HM004068, HM004069, HM004070	Chrysoviridae
Cryphonectria nitschkei chrysovirus 1	CnCV1	GQ290649, GQ290645, HM013825, HM013826	Chrysoviridae
Amasya cherry disease associated chrysovirus	AcdaCV	AJ781166, AJ781165, AJ781164, AJ781163	Chrysoviridae
Helminthosporium victoriae 145S virus	HvV145S	AF297176, AF297177, AF297178, AF297179	Chrysoviridae
Isaria javanica chrysovirus 1	IjCV1	KX898416, KX898417, KX898418, KX898419	Chrysoviridae
Macrophomina phaseolina chrysovirus 1	MpCV1	KP900886, KP900887, KP900888, KP900889	Chrysoviridae
Colletotrichum gloeosporioides chrysovirus 1	CgCV1	KT581957, KT581958, KT581959	Chrysoviridae
Wuhan insect virus 26	WIV26	KX882983.1	Totiviridae
Xanthophyllomyces dendrorhous virus L1A	XdVL1A	JN997472.1	Totiviridae
Xanthophyllomyces dendrorhous virus L1b	XdVL1b	JN997473.1	Totiviridae
Red clover powdery mildew-associated totivirus 3	RcpmaTV3	LC075488.1	Totiviridae
Tuber aestivum virus 1	TaV1	HQ158596.1	Totiviridae
Saccharomyces cerevisiae virus L-A	ScLAV	J04692.1	Totiviridae
Saccharomyces kudriavzevii virus L-A1	SKVLA1	KX601068.1	Totiviridae
Botryotinia fuckeliana totivirus 1	BITV1	AM491608.1	Victoriviridae
Magnaporthe oryzae virus 2	MoV2	AB300379.1	Victoriviridae
Alternaria arborescens victorivirus 1	AaVV1	LC086813.1	Victoriviridae
Fusarium poae victorivirus 1	FpVV1	LC150610.1	Victoriviridae
Gremmeniella abietina RNA virus L2	GaRV-L2	AY615210.1	Victoriviridae
Magnaporthe oryzae virus 3	MoV3	KP893140.1	Victoriviridae
Nigrospora oryzae victorivirus 1	NoVV1	KT428155.1	Victoriviridae
Penicillium aurantiogriseum totivirus 1	PaTV1	KT592305.1	Victoriviridae
Penicillium digitatum virus 1	PdTV1	KU933932.1	Victoriviridae
Phomopsis vexans RNA virus	PvRV	KP090346.1	Victoriviridae
Sclerotinia nivalis victorivirus 1	SnVV1	KT365894.1	Victoriviridae
Ustilagoidea vires RNA virus 3	UvV3	KF791042.1	Victoriviridae
Ustilagoidea vires RNA virus 5	UvV5	KT188753.1	Victoriviridae
Aspergillus foetidus slow virus 1	AfSV1	HE588147.1	Victoriviridae
Beauveria bassiana victorivirus 1	BbVV1	HE572591.1	Victoriviridae
Coniothyrium minitans RNA virus	CmRV	AF527633.1	Victoriviridae
Epichloe festucae virus 1	EfV1	AM261427.1	Victoriviridae
Gremmeniella abietina RNA virus L1	GaRV-L1	AF337175.1	Victoriviridae
Helicobasidium mompa totivirus 1-17	HmTV1-17	AB085814.1	Victoriviridae
Helminthosporium victoriae virus 190S	HvV190S	U41345.2	Victoriviridae
Magnaporthe oryzae virus 1	MoV1	AB176964.1	Victoriviridae
Rosellinia necatrix victorivirus 1	RnVV1	AB742454.1	Victoriviridae
Sphaeropsis sapinea RNA virus 1	SsRV1	AF038665.1	Victoriviridae
Sphaeropsis sapinea RNA virus 2	SsRV2	AF039080.1	Victoriviridae
Tolypladadium cylindrosporium virus 1	TcV1	FR750562.1	Victoriviridae
Beet cryptic virus 1	BcV1	ACA81389.1	Alphapartitivirus
Carrot cryptic virus	CcV	ACL93278.1	Alphapartitivirus
Amasya cherry disease-associated mycovirus	AcdaV	CAG77604.1	Alphapartitivirus
Cherry chlorotic rusty spot associated partitivirus	CcspaPV	CAH03668.1	Alphapartitivirus
Chondrostereum purpureum cryptic virus 1	CpcV1	CAQ53729.1	Alphapartitivirus
Flammulina velutipes browning virus	FvbV	BAH56481.1	Alphapartitivirus
Heterobasidion (annosum) partitivirus 12	HaPV12	AHL25151.1	Alphapartitivirus
Heterobasidion (annosum) partitivirus 13	HaPV13	AHL25153.1	Alphapartitivirus
Heterobasidion (parviporum) partitivirus 15	HpPV15	AHL25162.1	Alphapartitivirus
Heterobasidion (ecrustosum) partitivirus 3	HePV3	ACO37245.1	Alphapartitivirus
Rosellinia necatrix partitivirus 2	RnPV2	BAM78602.1	Alphapartitivirus
Podosphaera prunicola partitivirus 3	PpPV3	ATS94413.1	Alphapartitivirus
Vicia cryptic virus	VcV	AA39023.1	Alphapartitivirus
White clover cryptic virus 1	WccV1	AAU14888.1	Alphapartitivirus
Raphanus sativus cryptic virus 1	RscV1	AA51289.2	Alphapartitivirus
Diuris pendunculata cryptic virus	DpcV	AFQ95555.1	Alphapartitivirus
Dill cryptic virus 1	DcV1	AGY36136.1	Alphapartitivirus
Red clover cryptic virus 1	RccV1	AGY36138.1	Alphapartitivirus
Rhizoctonia solani dsRNA virus 2	Rsv2	AGY54938.1	Alphapartitivirus
Arabidopsis halleri partitivirus 1	AhPV1	BAV56959.1	Alphapartitivirus
Rose partitivirus	RPV	ANQ45203.1	Alphapartitivirus
Rosellinia necatrix partitivirus 7	RnPV7	BAT32942.1	Alphapartitivirus
Atkinsonella hypoxylon virus	AhV	AAA61829.1	Betapartitivirus

Virus Name	Virus Abbrev.	NCBI ID	Used for analysis of:
Cannabis (sativa) cryptic virus	CscV	AET80948.1	Betapartitivirus
Ceratocystis resinifera virus 1	CrV1	AAU26069.1	Betapartitivirus
Crimson clover cryptic virus 2	CccV2	AGJ83769.1	Betapartitivirus
Cucurbitaria piceae virus 1	CpV1	ALT08065.1	Betapartitivirus
Dill cryptic virus 2	DcV2	AGJ83771.1	Betapartitivirus
Fusarium poae virus 1-A11	FpV1-A11	AAC98734.1	Betapartitivirus
Fusarium poae virus 1-240374	FpV1-240374	YP_009272951.1	Betapartitivirus
Heterobasidion (parviporum) partitivirus 2	HpPV2	ADL66905.1	Betapartitivirus
Heterobasidion (parviporum) partitivirus 7	HpPV7	AEX87907.1	Betapartitivirus
Heterobasidion (irregulare) partitivirus 8	HiPV8	AFW17810.1	Betapartitivirus
Heterobasidion annosum P-type partitivirus	HaPPV	AAL79540.1	Betapartitivirus
Hop trefoil cryptic virus 2	HtcV2	AGJ83767.1	Betapartitivirus
Lentinula edodes partitivirus 1	LePV1	AQS27950.1	Betapartitivirus
Pleurotus ostreatus virus 1	PoV1	AAT07072.1	Betapartitivirus
Primula malacoides virus China/Mar2007	PmV2007	ABW82141.1	Betapartitivirus
Podosphaera prunicola partitivirus 4	PpPV4	ATS94415.1	Betapartitivirus
Red clover cryptic virus 2	RccV2	AGJ83765.1	Betapartitivirus
Rhizoctonia solani virus 717	RsV717	AAF22160.1	Betapartitivirus
Rosellinia necatrix partitivirus 1-W8	RnPV1-W8	BAD98237.1	Betapartitivirus
White clover cryptic virus 2	WccV2	AGJ83763.1	Betapartitivirus
Sclerotinia sclerotiorum partitivirus 1	SsPV1	AFR78160.1	Betapartitivirus
Rosellinia necatrix partitivirus 6	RnPV6	BAT24479.1	Betapartitivirus
Fusarium poae partitivirus 2	FpPV2	BAV56299.1	Betapartitivirus
Alternaria alternata partitivirus 1	AaPV1	KY352402	Zetapartitivirus
Botryosphaeria dothidea partitivirus 1	BdPV1	AGZ84316.1	Zetapartitivirus
Aspergillus ochraceous virus	AoV	ABV30675.1	Gammapartitivirus
Beauveria bassiana partitivirus 1	BbPV1	CUS18591.1	Gammapartitivirus
Botryotinia fuckeliana partitivirus 1	BfPV1	YP_001686789.1	Gammapartitivirus
Discula destructiva virus 1	DdV1	AAG59816.1	Gammapartitivirus
Discula destructiva virus 2	DdV2	AAK59379.1	Gammapartitivirus
Fusarium solani virus 1	FsV1	BAA09520.1	Gammapartitivirus
Gremmeniella abietina RNA virus MS1	GaVMS1	AAM12240.1	Gammapartitivirus
Ophiostoma (himal-ulmi) partitivirus 1	OhuPV1	CAJ31886.1	Gammapartitivirus
Penicillium stoloniferum virus F	PsVF	AAU95758.1	Gammapartitivirus
Penicillium stoloniferum virus S	PsVS	AAN86834.2	Gammapartitivirus
Pseudogymnoascus destructans partitivirus-pa	PdPVpa	YP_009259752.1	Gammapartitivirus
Ustilagoidea vires partitivirus	UvPV	AGO04402.1	Gammapartitivirus
Aspergillus fumigatus partitivirus-1	AfPV1	CAY25801.2	Gammapartitivirus
Colletotrichum acutatum RNA virus 1	CaRV1	AGL42312.1	Gammapartitivirus
Ustilagoidea vires partitivirus 2	UvPV2	AGR45851.1	Gammapartitivirus
Verticillium dahliae partitivirus 1	VdPV1	AGI52210.1	Gammapartitivirus
Penicillium aurantiogriseum partitivirus 1	PaPV1	ALO50131.1	Gammapartitivirus
Penicillium aurantiogriseum partiti-like virus	PaPIV1	YP_009182157	Epsilonpartitivirus
Rhizoctonia solani dsRNA virus 5	RsRV5	AVP26802	Epsilonpartitivirus
Beihai bamacle virus 12	BbV12	YP_009333370	Epsilonpartitivirus
Wuhan cricket virus 2	WcV2	YP_009345133	Epsilonpartitivirus
Hubei partiti-like virus 11	HplV11	YP_009329874	Epsilonpartitivirus
Rhizoctonia solani endomavirus 2	RsEV2	KT823701.1	Alphaendovirus
Ceratobasidium endomavirus A	CEVA	KX355142.1	Alphaendovirus
Alternaria brassicicola betaendovirus 1	AbEV1	KP239989.1	Betaendovirus
Gremmeniella abietina betaendovirus 1	GaRVXL1	DQ399289.1	Betaendovirus
Sclerotinia sclerotiorum endomavirus 2	SsEV2	KU299046.1	Betaendovirus
Tuber aestivum betaendovirus	TaEV1	HQ380014.1	Betaendovirus
Botrytis cinerea endomavirus 1	BcEV1	KU923747.1	Betaendovirus
Rosellinia necatrix endomavirus 1	RnEV1	LC076696.1	Betaendovirus
Sclerotinia sclerotiorum endomavirus 1	SsEV1	NC_021706.1	Betaendovirus
Diaporthe ambigua RNA virus 1	DaRV1	NP_037581	Ambiguiviridae
Magnaporthe oryzae RNA virus	MoRV	YP_009115495	Ambiguiviridae
Soybean leaf-associated ssRNA virus 1	SlassV1	ALM62232	Ambiguiviridae
Soybean leaf-associated ssRNA virus 2	SlassV2	ALM62236	Ambiguiviridae
Verticillium dahliae RNA virus	VdRV1	KV563741	Ambiguiviridae
Sclerotinia sclerotiorum umbra-like virus 1	SsUIV1	YP_009253998	Ambiguiviridae
Olive mild mosaic virus; gp1 == fusion (HP-RdRP)	OmMV	ANC98943.1	Ambiguiviridae
Tobacco necrosis virus A	TnVA	ASV49764.1	Ambiguiviridae
Olive latent virus 1	OIV1	AAZ43261.1	Ambiguiviridae
Maize chlorotic mottle virus	McMV	ATY93269.1	Ambiguiviridae
Rosa rugosa leaf distortion virus	RrdV	YP_007501034.1	Ambiguiviridae
Pelargonium flower break virus	PfbV	ABB79925.1	Ambiguiviridae
Hubei tombus-like virus 12	HtIV12	YP_009336735.1	Ambiguiviridae
Beihai tombus-like virus 8	BtIV8	NC_032951.1	Ambiguiviridae
Wenling tombus-like virus 1	WtIV1	NC_033043.1	Ambiguiviridae
Beihai tombus-like virus 7	BtIV7	YP_009337688.1	Ambiguiviridae
Beihai tombus-like virus 5	BtIV5	NC_033250.1	Ambiguiviridae
Aspergillus foetidus slow virus 2	AfSV2	CCD33025	Yadokarviridae
Penicillium aurantiogriseum foetidus-like virus	PaFIV	YP_009182156	Yadokarviridae

Table S2 continued

Virus Name	Virus Abbrev.	NCBI ID	Used for analysis of:
Yado-kari virus 1 (Rosellinia necatrix mycovirus 2-W1032/S6)	YkV1	BAT50982	Yadokarviridae
Fusarium poae mycovirus 2	FpMV2	YP_009272910	Yadokarviridae
Pig calicivirus NC-USA	PcV	ADG27878	Yadokarviridae
Norwalk-like virus	NiV	ADF45588	Yadokarviridae
Broad bean wilt virus 2	BbWV2	BAB18312	Yadokarviridae
Rabbit hemorrhagic disease virus	RHDV	AAA47285	Yadokarviridae
Sapovirus (Hu/Dresden/pJG-Sap01/DE)	SV-Hu	AAU09265	Yadokarviridae
Phomopsis longicolla RNA virus 1	PIRV1	YP_009345044.1	Oourmia-like virus
Botrytis ourmia-like virus HAZ2-3	BOiV	YP_009182165.1	Oourmia-like virus
Soybean leaf-associated ourmiavirus 1	SlaOV1	ALM62238.1	Oourmia-like virus
Magnaporthe oryzae ourmia-like virus	MoOV	SBQ28480.1	Oourmia-like virus
Cassava virus C	CvC	YP_003104770.1	Oourmia-like virus
Epirus cherry virus	EcV	YP_002019754.1	Oourmia-like virus
Oourmia melon virus	OmV	YP_002019757.1	Oourmia-like virus
Grapevine associated namavirus-1 HAZ1-4	GaNV1	YP_009182162.1	Mitovirus
Cryphonectria parasitica mitovirus 1-NB631	CpMV1-NB631	NP_660174.1	Mitovirus
Buergenerula spartinae mitovirus 1	BsMV1	AHY03257.1	Mitovirus
Sclerotinia sclerotiorum mitovirus 7	SsMV7	AHF48623.1	Mitovirus
Rhizoctonia solani mitovirus 11	RsMV11	ALD89116.1	Mitovirus
Rhizoctonia solani mitovirus 2	RsMV2	ALD89121.1	Mitovirus
Botrytis cinerea mitovirus 2 HAZ1-2	BcMV2	NC_028471.1	Mitovirus
Sclerotinia sclerotiorum mitovirus 5	SsMV5	KJ462510.1	Mitovirus
Sclerotinia sclerotiorum mitovirus 1	SsMV1	JQ013377.1	Mitovirus
Ophiostoma mitovirus 5	OnuMV5	NC_004053.1	Mitovirus
Neofusicoccum luteum mitovirus 1	NIMV1	NC_035114.1	Mitovirus
Loramycetes juncicola mitovirus 1	1LjMV1	PRJNA372853	Mitovirus
Sclerotinia sclerotiorum mitovirus 6	SsMV6	KF913881.1	Mitovirus
Cryphonectria cubensis mitovirus 1a	CcMV1a	AY328476.1	Mitovirus
Hymenoscyphus fraxineus mitovirus 1	HfMV1	KT809403.1	Mitovirus
Sclerotinia sclerotiorum mitovirus 15	SsMV15	KF913890.1	Mitovirus
Sclerotinia sclerotiorum mitovirus 9	SsMV9	KF913884.1	Mitovirus
Cryphonectria hypovirus 1	CHV1	ATZ76109.1	Alphahypovirus
Cryphonectria hypovirus 2	CHV2	NP_613266.1	Alphahypovirus
Fusarium graminearum hypovirus 1	FgHV1	YP_009011065.1	Alphahypovirus
Rosellinia necatrix hypovirus 2	RnHV2	BBB86794.1	Alphahypovirus
Fusarium langsethiae hypovirus 1	FlHV1	YP_009330037.1	Alphahypovirus
Fusarium graminearum hypovirus 2	FgHV2	YP_009130646.1	Alphahypovirus
Fusarium poae hypovirus 1	FpHV1	BAV56305.1	Alphahypovirus
Rosellinia necatrix hypovirus 1	RnHV1	YP_009448196.1	Alphahypovirus
Cryphonectria hypovirus 4	CHV4	YP_138519.1	Betahypovirus
Valsa ceratospema hypovirus 1	VcHV1	YP_005476604.1	Betahypovirus
Phomopsis longicolla hypovirus	PIHV	YP_009051683.1	Betahypovirus
Cryphonectria hypovirus 3	CHV3	NP_051710.1	Betahypovirus
Sclerotinia sclerotiorum hypovirus 1	SsHV1	YP_004782527.1	Betahypovirus
Fusarium graminearum dsRNA mycovirus-1	FgV1	NC_006937	Fusariviridae
Fusarium poae fusarivirus 1	FpFV1	NC_030868	Fusariviridae
Macrophomina phaseolina single-stranded RNA virus 1	MpssV1	KP900890	Fusariviridae
Neofusicoccum luteum fusarivirus 1	NfFV1	ARO52688	Fusariviridae
Penicillium aurantiogriseum fusarivirus 1	PaFV1	NC_028467	Fusariviridae
Penicillium roqueforti ssRNA mycovirus 1	PrssV1	NC_024699	Fusariviridae
Pleospora typhicola fusarivirus 1	PfFV1	NC_028470	Fusariviridae
Rosellinia necatrix fusarivirus 1	RnFV1	NC_024485	Fusariviridae
Sodiomyces alkalinus fusarivirus 1	SaFV1	APC23179	Fusariviridae
Alternaria brassicicola fusarivirus 1	AbFV1	NC_029056	Fusariviridae
Nigrospora oryzae fusarivirus 1	NoFV1	NC_031960	Fusariviridae
Sclerotinia sclerotiorum fusarivirus 1	SsFV1	NC_027208	Fusariviridae
Macrophomina phaseolina tobamo-like virus 1	MpTIV1	NC_025674.1	Tobamo-like
Podosphaera prunicola tobamo-like virus 1	PpTIV1	KY420046.1	Tobamo-like
Sclerotinia sclerotiorum RNA virus L	SsRVL	ACE88957.1	MiRV1
Rhizoctonia solani RNA virus 1	RsRV1	ANR02702	MiRV1
Rhizoctonia solani RNA virus 2	RsRV2	ANR02701	MiRV1
Rhizoctonia solani RNA virus 3	RsRV3	ANR02700	MiRV1
Potato Virus Y	PVY	ARI46497.1	Outgroup (various)
Tobacco Mosaic Virus	TMV	NP_597747.1	Outgroup (various)
Simian Rotavirus A	SRVA	P22678.1	Outgroup (various)