Large DNA viruses in early diverging fungal genomes are relics of past and present infections

Keywords: giant virus, early-diverging fungi, endogenous viral elements (EVEs), NCLDV, Nucleocytoviricota, mycovirus

Jillian M. Myers[1], Frederik Schulz[2], Saleh Rahimlou[1], Kevin R. Amses[1◊], D. Rabern Simmons[1◊], Jason E. Stajich[3,4✉], Timothy Y. James[1✉]

Author Affiliations:
[1]Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI USA
[2]DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[3]Microbiology and Plant Pathology, University of California Riverside, Riverside, CA USA
[4]Institute for Integrative Genome Biology, University of California Riverside, Riverside, CA USA
◊current affiliations: University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA (KRA), Botany and Plant Pathology, Purdue University, West Lafayette, Indiana, USA (DRS)
✉email: tyjames@umich.edu; jason.stajich@ucr.edu

Abstract

Giant viruses of the phylum Nucleocytoviricota have captured researchers' attention due to their increasingly recognized impacts on eukaryotic genome evolution. Their origins are hypothesized to predate or coincide with the diversification of eukaryotes and their natural hosts span the eukaryotic tree of life. But surprisingly, giant viruses have not been found in the fungal kingdom, though recent metagenomic work suggests a putative association. Here we show that early-diverging fungal lineages maintain both "viral fossils" and host active infections by giant viruses, which form a monophyletic clade that we name *Mycodnaviridae*. Viral genomes described here span up to 350kb and encode over 300 genes, including many genes characteristic of giant viruses. Interestingly, we observed variation in infection status among the isolates including apparent active infection and transcriptionally-suppressed states, suggestive that viral activation may be constrained to certain life stages of the host. Our experimental findings add to the scant few natural virus-host systems available in culture for the study of giant viruses. These viruses have likely shaped the early evolution of these fungal lineages and should prove a useful model for their study.

Main

In the entire Fungal Kingdom, only a handful of DNA viruses are known–all small circular single-stranded (CRESS) viruses (Li et al. 2020; Yu et al. 2010). This apparent rarity of DNA viruses in

fungi likely results from taxonomic biases in searches, rather than a true biological phenomenon. For example, we previously demonstrated that by expanding the taxonomic scope of considered fungal hosts to include early-diverging fungal phyla (EDF), the known RNA mycoviral diversity was greatly increased (Myers et al. 2020). EDF include many lineages outside of Dikarya—the group containing the majority of described fungal taxa and viral hosts; the current fungal taxonomic division includes 8 phyla of zoosporic fungi that reproduce with motile spores and represent a particularly understudied paraphyletic assemblage near the base of the fungal phylogeny (James et al. 2020). The presence of DNA viruses in EDF has been almost entirely unexplored (but see Clemons et al. 2023), and indeed, the finding of large DNA viruses in basal lineages of fungi, which we report here, was a serendipitous discovery resulting from a large genomic sequencing effort across early diverging fungal phyla (Amses et al. 2022).

Recent studies have identified relics of DNA virus elements in fungal genomes or found evidence of associations between DNA viruses and fungi in metagenomes. The first report came after mining the eukaryotic genomes that pointed toward the presence of endogenized viral genes in *Gonapodya prolifera* and *Allomyces macrogynus* (Gallot-Lavallée and Blanc 2017). Further evidence was generated through the genome of sylvanvirus, a giant virus identified by soil metagenomics, which contained a high proportion of fungal homologs (Schulz et al. 2018). Shortly after, in a global metagenomics study of giant viruses and hosts, horizontal gene transfers between multiple groups of fungi and multiple previously described giant virus lineages and metagenomic-assembled-viral-lineages were proposed, suggesting that historical fungal-giant virus associations have spanned numerous fungal and giant virus lineages (Schulz et al. 2020). In a search that spanned the breadth of the fungal kingdom available at the time, Gong *et al.* (2020) found giant virus-like genes endogenous in the genomes of 15 species, which grouped into two distinct viral lineages based on the family B DNA polymerase; coevolutionary analysis suggested independent integration events. More recently, a 1.5Mb endogenous viral element (EVE), related to Asfarviridae, was described in the genome of *Rhizophagus irregularis*, an arbuscular mycorrhizal fungus (Zhao et al. 2023). Bhattacharjee (2023) found two Nucleocytoviricota genomes containing fungal genes and, because ascomycetous fungi are the dominant eukaryotes found in the subsurface oceanic crustal fluids from which these samples derived, suggested a novel fungus-giant virus specific association. Despite these tantalizing findings that suggest infection in fungi by DNA viruses may be more common than appreciated, novel clades that include active giant viruses in fungi have not been characterized.

Here we report both EVEs and active infection by large DNA viruses in EDF. The recovered viral genomes form a monophyletic clade within the Nucleocytoviricota related to orders Imitervirales and Algavirales, which we name family *Mycodnaviridae*. These findings expand the known host range of giant viruses to include all major eukaryotic lineages. DNA viruses have played, and continue to play, a major role in the evolution of EDF, as previously speculated (Medina 2016). We show that fungal giant viruses have a global distribution, which begs inquiry into the implications of these infections for host traits, host genome evolution, and the function of ecosystems.

Results

Giant virus genes and genomes in the early-diverging fungi

Using Hidden Markov Model (HMM) searches of the predicted proteins, we identified homologs of the giant virus major capsid protein (MCP; NCVOG0022) in the genomes of 17/141 fungi (Fig S1, Table S1). The distribution of MCP homologs is restricted to the EDF and, particularly, in zoosporic lineages; no MCPs were found by hmmseach in non-zoosporic lineages (n= 40). We further searched all fungal genomes in Mycocosm (Grigoriev et al. 2014; n = 2,543) for MCP homologs by blastp, which resulted in one hit to a non-zoosporic EDF species: *Dimargaris cristalligena* (phylum Zoopagomycota), a mycoparasite of other EDF. Among the zoosporic lineages, MCPs were found in Chytridiomycota (n=12), Monoblephidomycota (n=2), and Blastocladiomycota (n=3). Blastocladiomycota are a particularly poorly sampled group, and thus our genomic searches were limited (n=7). However, MCPs appear to be common in the most highly sampled genus within this phylum, *Allomyces.* Using available sequences, we designed primers targeting the MCP gene of *Allomyces* and found 30/58 isolates were positive for capsid by PCR.

We first attempted to characterize the contigs containing MCP genes. These contigs vary widely among the isolates in gene content, GC composition, intergenic distance, and other metrics, suggesting numerous independent endogenization events (Fig 1 A). In three isolates (*Podochytrium* sp. JEL797, *Gaertneriomyces semiglobifer* Barr43, and *Fimicolochytrium jonesii* JEL569), MCP was the only gene of viral provenance on contigs which otherwise contain primarily fungal genes and have GC content and intron density similar to other host contigs (Fig S2). It is possible that these species were historically infected with giant viruses and horizontal gene transfer events left viral genes scattered throughout the host genome. Or, perhaps a virus integrated into the host genome and, over time, underwent gene loss. The isolation of some giant virus orthologous genes (GVOGs) on small fungal contigs could, alternatively, be an artifact of variable genome quality; perhaps in some cases we have misidentified proviruses as small EVEs. Future long-read sequencing approaches may help to clarify this result, but, generally, there is much that remains to be understood about the nature and functionality of endogenized giant viruses.

In other isolates, contigs with MCPs appear as mosaics of virus, prokaryote, and fungal genes, which include multiple additional GVOGs, and have markedly greater gene density compared to host contigs, consistent with reference giant virus genomes (Fig 1 A, B; Fig S2). These giant virus-like contigs range up to nearly 350kb, and are suggestive of either very recent endogenization or proviral state. Interestingly, in five isolates the MCP contig contains an integrase-encoding gene without homology to the integrases commonly found in fungi. Indeed, one of these, the integrase in *Synchytrium microbalum,* was recently annotated as a MELD virus element (Starrett et al. 2021). The authors proposed the name “MELD” (midsize eukaryotic linear dsDNA) to describe the polyphyletic group of viruses which frequently engage in horizontal gene transfer and so are particularly challenging to pin down taxonomically. Our results suggest that these “apparently promiscuous” HGT events could also occur between

midsize and giant viruses. These integrases may also be a defining characteristic of the *Mycodnaviridae* clade (described below). Yet another possibility is the involvement of these integrases in a fungal anti-giant-virus defense system. In a recent landmark study, Fanzor(Fz) endonucleases were structurally characterized from an EDF, demonstrating the conservation of core regions of this protein and other RNA-guided endonucleases including Cas12 (Saito et al. 2023). This discovery of a programmable CRISPR-Cas-like system in eukaryotes is particularly exciting and relevant to the work we describe here, as Fz1 is particularly diverse in EDF and prominent in giant viruses (Saito et al. 2023). This connection is remarkable and a promising area of future study.

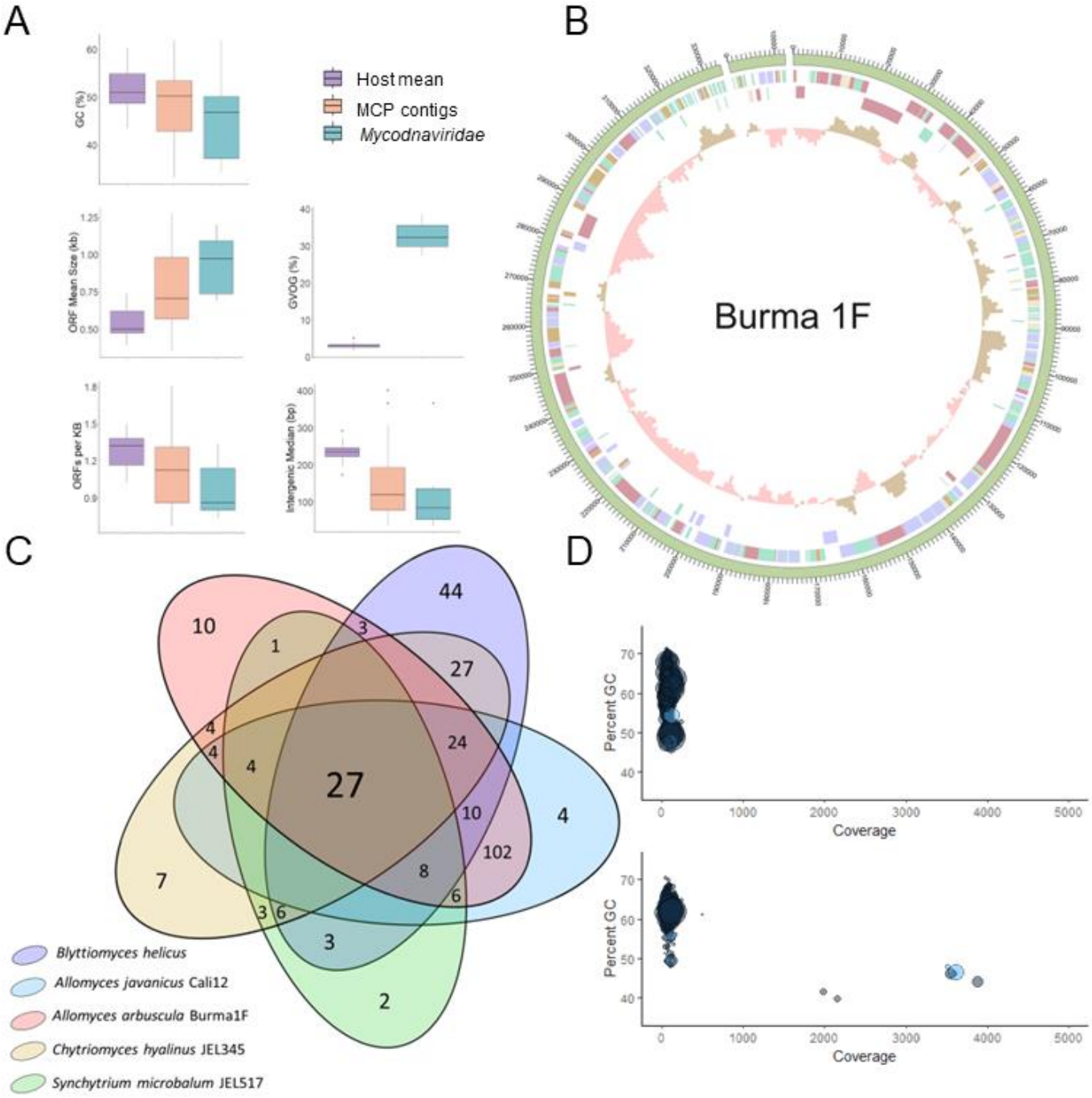

Fig 1. Genomic features of *Mycodnaviridae* (A) Boxplots of genome statistics for mean values across all contigs in all fungal genomes ("host mean", purple), all contigs with at least one major capsid protein (MCP) gene ("MCP contigs", pink), and MCP contigs that met the criteria set for orthology analysis (ViralRecall score >2, viral region length > 70%, intron density >25% different from host; "*Mycodnaviridae"*, teal). (B) The genome of Allomyces arbusculus Burma1F mycodnavirus 1 is a mosaic of bacteria; viral, and eukaryotic genes. Contigs are represented by the outermost ring in the circularized genome plot. From outside to inside: ORFs on the forward strand, ORFs on the reverse strand (both color coded by provenance), GC skew. (C) Venn diagram indicating the number of orthologous gene clusters shared between the 5 most complete fungal giant virus genomes (D) Plots of the GC content and coverage for two *Allomyces* sp. genome assemblies: *A. anomalus* BEA2 (upper) and *Allomyces* sp. Cali8 (lower). Contigs that include an MCP are shaded blue, size of data points reflect contig length.

The most contiguously assembled viral genomes were found from two strains of *Allomyces*. A. arbusculus Burma1F mycodnavirus 1 includes two contigs of total size 345,111 bp, encoding 314 genes with likely provenance, based on homology, assigned as follows: 10% fungi, 14% bacteria, 20% virus, 4% putatively *Mycodnaviridae*-specific, and 52% unknown/ORFan (Fig 1B). The isolate *A. javanicus* California12 has multiple large contigs (each approximately 80-288kb) with giant virus signatures (Fig S3). Interestingly, when binned by tetramer frequency, the viral contigs in *A. javanicus* California12 form two bins suggesting the presence of two distinct viruses. The two bins are roughly similar in size (212,591bp and 287,844 bp) and show some duplication of hallmark GVOGs: each bin contains copies of MCP, D5-like helicase-primase, DNA PolB, mRNA capping enzyme large subunit, and superfamily II helicase. The larger bin contains an additional five GVOGs (YqaJ recombinase, RNA ligase, disulfide (thiol) oxidoreductase, transcription initiation factor IIB, and A32-like packaging ATPase). *A. javanicus* strains are believed to be hybrids (Emerson 1941) and the two viruses could have originated from the two parents of this diploid sporophyte.

Mycodnaviridae is a monophyletic clade of fungal giant viruses

A phylogeny of representative giant viruses including the 7 fungal samples with at least 4/7 core genes (Aylward et al. 2021) show that fungal giant viruses form a monophyletic clade together with two giant virus metagenome-assembled genomes (GVMAGs) of the previously proposed Imitervirales family IM_02 and in a sister position to Imitervirales IM_18. Our phylogenetic analysis showed that less complex phylogenetic models (WAG) and reconstruction algorithms (FastTree, IQ-tree2 -fast) place the IM_02-IM_18 cluster basal to other Imitervirales. However, with more complex phylogenetic models (IQtree2 with LG+I+G4, LG4X) the IM_02-IM_18 cluster was split off from Imitervirales and groups basal to Pandoravirales and Algavirales. The monophyletic grouping of the fungal giant viruses with IM_02 was well supported across supermatrix-based species trees and single protein trees for each of the markers in the GVOG7 set of Nucleocytoviricota hallmark genes (Aylward et al. 2021). Both IM_02 and IM_18 previously consisted exclusively of GVMAGs but since the new genomes added can be connected with fungal hosts, we dub IM_02 *Mycodnaviridae* (Fig 2 A). Our analysis also suggests that IM_02 together with IM_18 is not part of the order Imitervirales but may represent a novel order in the Nucleocytoviricota. Further detailed phylogenomic analysis after improved taxon sampling will be required to provide additional evidence for such a relationship. The hosts of IM_02 and IM_18 are unknown, but the clustering of these clades with giant viruses here

described might suggest their fungal-association. Further, the fungal giant viruses we identify have worldwide provenance, suggesting global distribution of these GVMAGs (Fig 2 B).

We investigated whether a core suite of genes exist that functionally define the *Mycodnaviridae* by conducting orthology analysis on the most complete viral genomes–five have at least 9 of the 20 core GVOGs most likely to have been vertically inherited (Yutin et al. 2009). This resulted in 27 gene clusters shared by all five (Fig 1C) though this is likely an underestimate due to variation in genome completeness, suggested by variation in genome size and resulting from imperfect criteria for distinguishing viral from host contigs and/or sequencing quality. In addition to GVOGs such as MCPs, VLTF3, and YqaJ-like recombinase are gene orthogroups involved in DNA biosynthesis such as DNA helicase, ribonucleotide reductase, and proliferating cell nuclear antigen, which are also GVOGs found in other giant viruses. Other shared gene clusters include a cysteine protease, protein with N-acetylglucosaminyltransferase activity (possibly involved in cell wall modification, six genes of unknown function, and five novel genes apparently unique to these viruses.

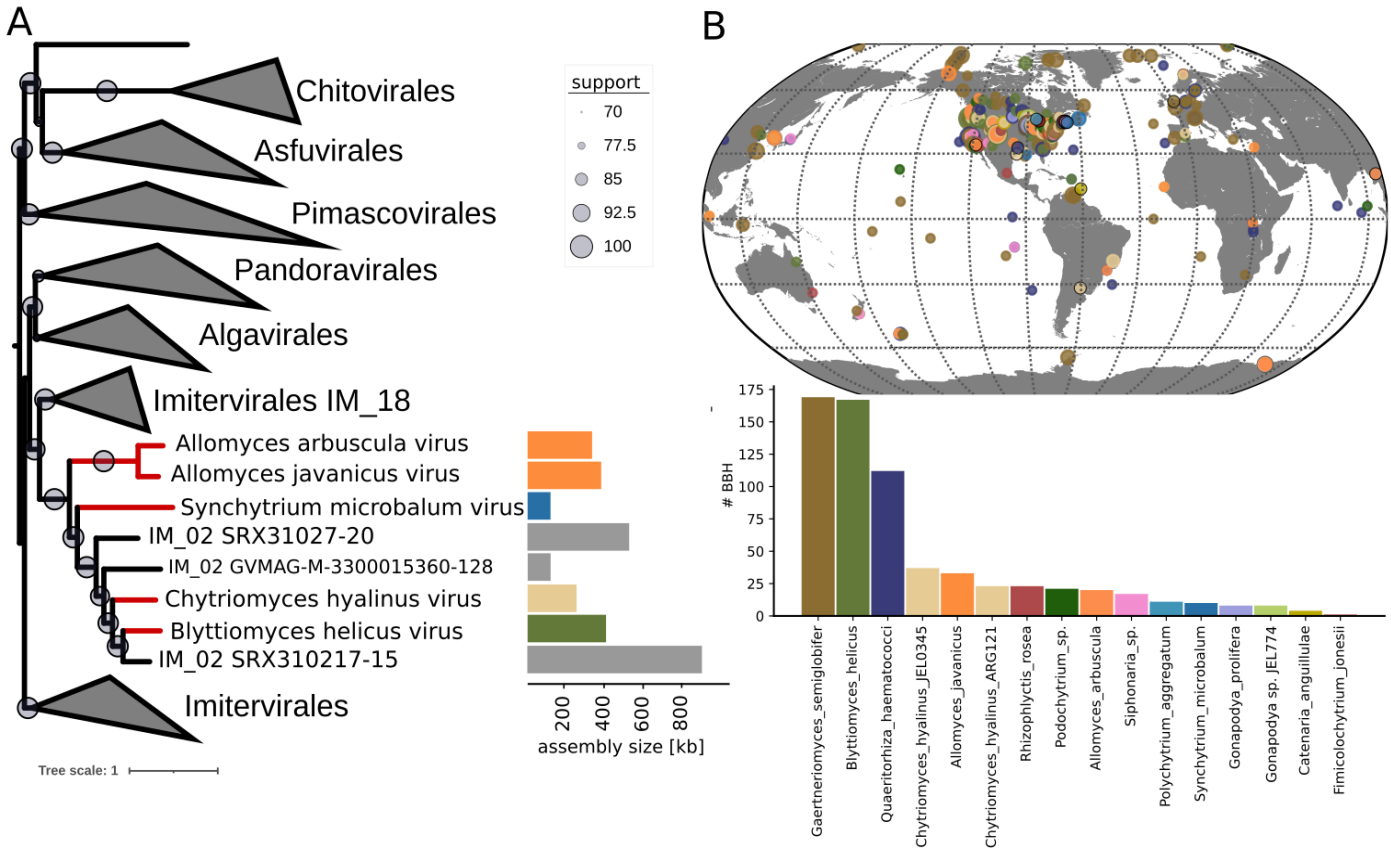

Fig 2. Phylogenetic position of the *Mycodnaviridae* and global distribution. A. Position of the *Mycodnaviridae* clade in the species tree of the Nucleocytoviricota. Red branches indicate *Mycodnaviridae* genomes. Tree is rooted at Pokkesviricetes. B. Global distribution of the *Mycodnaviridae* based on the major capsid protein. Colored circles correspond to metagenomic MCP sequences related to *Mycodnaviridae* identified in this study and circles with a black outline indicate geographic location of *Mycodnaviridae* and hosts described in this study.

Genome coverage, viral gene expression, and infected phenotype

One primary distinguishing feature between viruses and EVEs is read coverage of the viral contigs because sequencing during an active infection is expected to result in extremely high coverage of viral reads. We found no clear pattern between coverage of contigs containing MCPs and coverage of host contigs, with one exception. We sequenced an additional 9 strains of *Allomyces* that resulted in one additional isolate, *Allomyces* sp. Cali8, with 8 contigs (total 309,231 bp) averaging nearly 50-fold higher coverage than the rest of the genome (Fig 1D, lower plot). Of those, two contigs include MCPs and a combined 7 additional core GVOGs as well as a viral-like integrase. Overall, these results strongly indicate an active infection by a giant virus in this strain of *Allomyces.* Other sequenced *Allomyces* strains had contigs similarly rich with giant virus hallmark genes but lacking differences in coverage estimated by read depth. It is plausible that in these strains the conditions for triggering viral release were not met under laboratory culture conditions. We suspect that more of the *Allomyces* strains than suggested by the coverage data contain provirus that is perhaps capable of activating and forming an infectious propagule, though the triggering mechanism for viral release is unknown.

We investigated the expression of viral and fungal genes in two of the isolates using either zoospores, tissue containing mycelium and possibly zoosporangia and/or resting sporangia, or mycelium triggered to produce zoosporangia as starting material. We found almost no evidence of expression from the viral genes. Particularly, for strain Burma1F we found that genes in the large contig had very little expression except of the genes in the flanking region (Fig S4). These data suggest active silencing of the viral genes, and the expression helps define the border of the integrated virus.

In an initial attempt to identify functional impacts of giant virus presence, strains of *Allomyces* with and without detectable MCP genes were grown on agar and liquid Emerson's YpSS media to test if the presence of MCP influenced growth rate of the fungal strain. No significant differences in growth rate ($t(10) = -1.10$, $p = 0.30$) and biomass accumulation ($t(10) = -0.06$, $p = 0.95$) were observed (Fig S5, A & B). We also compared the thermal maximum of *Allomyces* strains with and without MCP grown on solid YpSS media. We found no significant difference ($t(10) = -0.25$, $p = 0.80$; Fig S5 C) depending on presence of MCP, however, one of the strains, Cali8, showed sectoring at 42 C. Interestingly, this is the strain that has evidence of active viral infection; however after additional subculturing the strain's phenotype became more similar to wild type. Overall, these results suggest that there is typically no effect of the endogenous virus on *Allomyces* growth insofar as we tested. Our finding of a lack of effect of the virus on fungal host phenotypes may be related to the absence of expression of virus genes when present in the fungal chromosome. If our hypothesis on the replication strategy of these viruses is correct, the endogenous viruses may be silenced in the host genome, a pattern consistent with transcriptional patterns of phaeoviruses integrated in their algal hosts (Cock et al. 2010).

Transmission electron microscopy

The three phyla of fungi in which we found evidence of giant viruses have life cycles similar to some algae, which may facilitate infection by viruses with persistent infection strategies, such as

the phaeoviruses in Algavirales. The close phylogenetic relationship between Algavirales and the group of viruses we describe here adds weight to this hypothesis. In both algae and zoosporic fungi there exist life-stages (spores and gametes) consisting of flagellated single-cells lacking a protective cell wall. The phaeoviruses have evolved stable coexistence with their hosts through a persistent viral strategy whereby spores or gametes are infected by virus particle(s) that then integrate into the host genome (Müller and Parodi 1993). Development proceeds with one copy of the virus integrated into the DNA of each cell (Meints et al. 2008). Thalli often appear symptom-free until some unknown mechanism triggers virus emergence from the genome. Reproductive organs (sporangia) may develop irregularly as they become factories for viral replication (Henry and Meints 1992; Stevens et al. 2014). We hypothesize that fungal giant viruses in zoosporic lineages may share similar life cycle characteristics. Transmission electron microscopy of *A. javanicus* California12 zoospores shows cytoplasmic virus-like particles (VLPs) (Fig 3). Similarly to phaeoviruses, the VLPs are round to hexagonal, 150-175 nm diameter with an approximately 100 nm granular core. VLPs appear in low abundance in zoospores, consistent with expectations that zoospores are recipients of infective viruses that migrate to the nucleus and, upon zoospore encystment, integrate into the host genome. We report this finding with caution, however, as VLPs have not yet been isolated and directly sequenced.

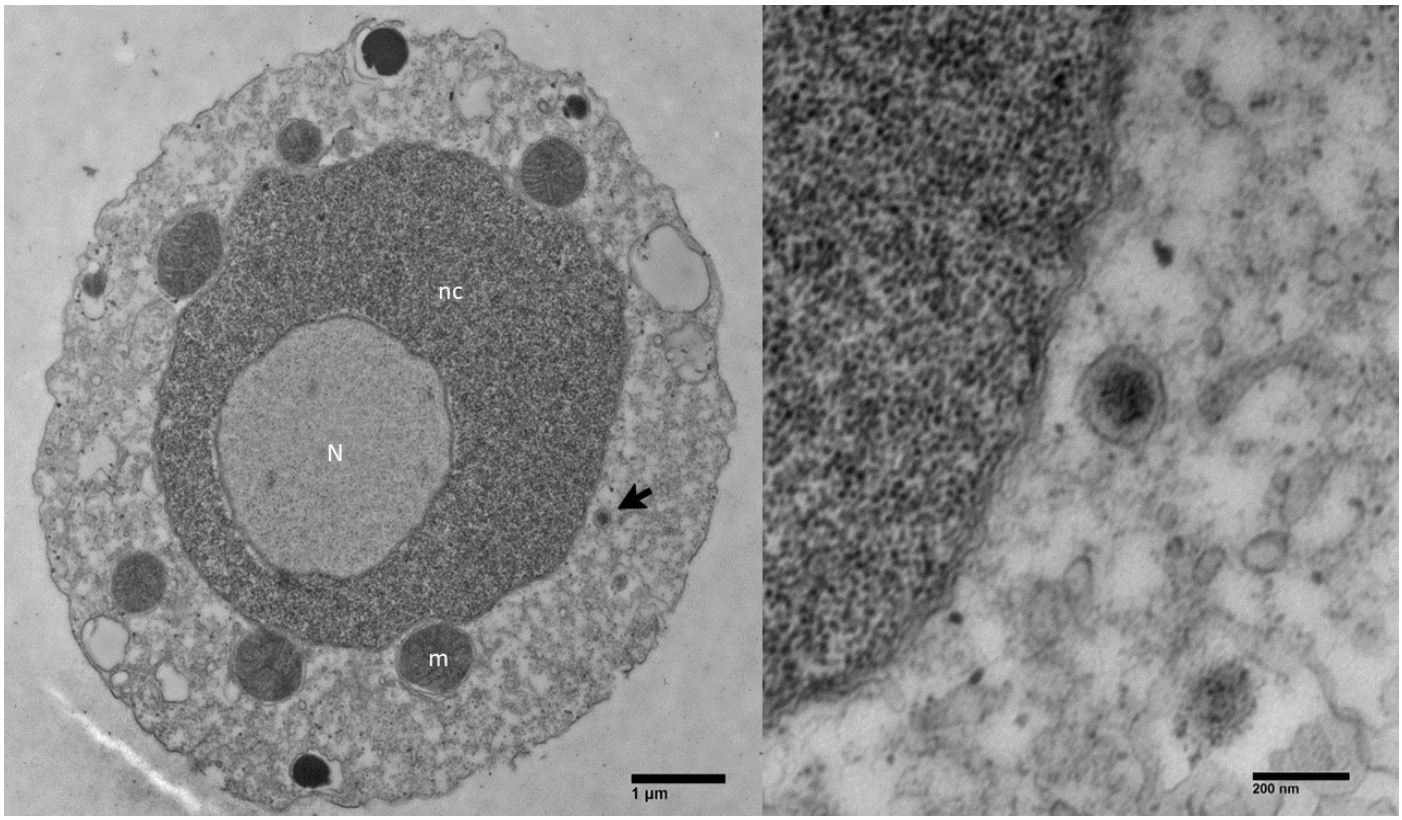

Fig 3. Transmission electron microscopy imaging of *A. javanicus* California12 zoospore cross-section at 2500x (left) containing a VLP (arrowhead) near the nuclear cap (nc), and the VLP at 15000x magnification (right) . N=nucleus, m= mitochondrion.

Discussion

We report here our finding of giant virus infection in fungi–both as endogenous elements altering fungal genome structure and as active viruses. Unlike the highly fragmented and duplicated giant virus relics littered throughout algal genomes (Moniruzzaman et al. 2020), these endogenous elements can appear as intact genomes up to 345 kb in length, suggestive of very recent endogenization events or current infection as stable provirus. Others appear to be relics of past infections but genome coverage data in one isolate of *Allomyces* suggests an active infection. Electron micrographs of putatively infected zoospores show virus-like particles (VLPs). Based on the exclusivity of this giant virus group to early-diverging fungi, phylogenetic placement of *Mycodnaviridae*, presence of integrase genes, morphology of VLPs, and read coverage of viral sequences, we speculate that these fungal giant viruses may adopt a non-lethal, persistent provirus replication strategy intimately linked to host life cycle, similar to the phaeoviruses in Algavirales. The apparent restriction of *Mycodnaviridae* to zoosporic lineages is consistent with this hypothesis, as is transcriptome profiling, though a mechanism triggering viral release remains unknown.

Notably, the most contiguous viral genomes were found in *Allomyces* spp., from cultures that were recovered from filter paper stocks after 30 years of storage. This fact demonstrates desirable aspects of this fungus that bolster motivation for its revival as a model system. *Allomyces* has historically, and currently though with much less frequency, been a model for studying biological phenomena such as anisogamy (Olsen 1984; Phadke 2018), sensorimotor systems including phototaxis and chemotaxis (Machlis 1969; Swafford and Oakley 2018), and even cellular differentiation (Cantino and Lovett 1964). The development of *Allomyces* as a model for studying giant viruses has enormous potential. Much of what is known about giant virus biology has been learned from protist-virus interactions or metagenomics where the host is unknown. The development of additional laboratory model systems of giant viruses and their natural hosts could advance our understanding of virus-host interactions, including immune response, viral “reprogramming” of host cells, metabolic impacts, and their ecosystem-level consequences.

Like all cellular organisms, the evolution of the fungi has been impacted by interactions with viruses. Giant viruses are hypothesized to have originated with the diversification of eukaryotes, and metagenomic sequence data demonstrates their abundance in the environment. The lack of evidence of active infection by giant viruses in fungi is therefore surprising, though previous reports have shown their association (Zhao et al. 2023; Bhattacharjee et al. 2023; Gong 2020; Schulz et al. 2020; 2018). Most interestingly, the EVEs reported by others are of distant relation to the clade of giant fungal viruses we report here. The 1.5Mb EVE reported from EDF *Rhizophagus irregularis* (phylum Mucoromycota) was found to be most closely related to viruses in the family *Asfarviridae* in order Asfuvirales (Zhao et al. 2023), which is sister to *Poxviridae*, in a basal position in the Nucleocytoviricota (Aylward et al. 2021). The virus genomes putatively associated with Ascomycota in subsurface oceanic crust cluster with the Mesomimiviridae; they are thus more closely related to *Mycodnaviridae* than the EVE in *R. irregularis* but are still distinct. Lastly, the EVEs reported by Gong *et al.* (2020) are most closely related to a *Pithovirus*

sibericum, in yet another giant virus order, the Pimascovirales. Taken together, these studies suggest that fungi have historically and, perhaps, currently been host to multiple distinct lineages of giant viruses.

Given this evidence for possible ancient association, the implications to fungal evolution are intriguing to consider since viral gene acquisition can be an impetus for evolutionary innovation (Frank and Feschotte 2017). In the Fungi, a transcription factor involved in cell-cycling that is, surprisingly, unique to fungi is hypothesized to have been acquired by the fungal ancestor through lateral gene transfer from a virus (Medina et al. 2016). Phylogenetic evidence points to the source as either a giant virus or phage-infected bacterium. We searched MCP-containing contigs but found no homologs containing the diagnostic viral-like KilA-N domain of this cell-cycle protein, though this does not preclude giant viruses as the original source. Other evolutionary innovations resulting from giant virus association could involve novel metabolic functions. Especially in GVMAGs, a vast diversity of metabolic genes has been found (Schulz et al. 2020, Moniruzzaman et al. 2020) including many involved in carbon and nitrogen metabolism. Assessing the genetic and functional impact of giant virus infection on fungal metabolism should be a future research priority. To understand the impact of the viruses, we need to resolve when and if the viruses are expressed during the life cycle.

The findings we report here inspire numerous questions that beg further investigation. How is viral expression and virion production related to the host life cycle, and vice-versa? What are the metabolic impacts of infection, and how important are they on the global scale? What defense systems do fungi employ against giant viruses? To what extent have giant viruses shaped genome evolution in fungi? Giant virus model systems in zoosporic fungi may be the key to exposing these latent revelations.

Methods

Starts page 15, below.

Data Availability

Mycodnaviridae genome assemblies and alignment for the Nucleocytoviricota phylogeny were deposited in Deep Blue Data, the University of Michigan's institutional data repository (DOI: https://doi.org/10.7302/p5ay-ng61).

Acknowledgements

Cultures and DNAs were obtained from the Collection of Zoosporic Eufungi at the University of Michigan. Next-generation sequencing was performed at the Advanced Genomics Core at the University of Michigan. Transmission electron microscopy was conducted at the University of Michigan's Department of Molecular, Cell, and Developmental Biology Imaging Core; we heartily thank Gregg Sobocinski for his technical assistance with sectioning and training on the Jeol. We thank Mia Sinks for technical assistance with *Allomyces* cultures. This project was

funded in part by the National Science Foundation, grant DEB-1929738 to TYJ, DEB-1557110 to JES. Data analyses performed at the High-Performance Computing Cluster at the University of California Riverside in the Institute of Integrative Genome Biology were supported by NSF grants DBI-1429826, DBI- 2215705, and NIH grant S10-OD016290. KRA was supported by the following National Institutes of Health training grant: “Michigan Predoctoral Training in Genetics” [T32GM007544]. JES and TYJ are fellows of the CIFAR program Fungal Kingdom: Threats & Opportunities. The work conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.

Author Contributions

JES, TYJ, and JMM conceived of and planned the study. DNA and RNA preparation, PCR, transcriptome assembly, viral content identification, and orthology analysis was conducted by JMM. Giant virus phylogeny and distribution analyses were performed by FS. Genomes assembly and annotation was conducted by JMM, KRA, and SR. Bioinformatic assistance was provided by KRA. Transcriptome analysis was conducted by SR. Summary statistics were conducted by JES. Growth experiments were performed by TYJ. Transmission electron microscopy was performed by JMM and DRS. Manuscript was written by JMM with contributions from TYJ and FS. All authors read and approved the final manuscript.

References

Amses, Kevin R., D. Rabern Simmons, Joyce E. Longcore, Stephen J. Mondo, Kensuke Seto, Gustavo H. Jerônimo, Anne E. Bonds, et al. 2022. “Diploid-Dominant Life Cycles Characterize the Early Evolution of Fungi.” *Proceedings of the National Academy of Sciences* 119 (36): e2116841119. https://doi.org/10.1073/pnas.2116841119.

Aylward, Frank O., Mohammad Moniruzzaman, Anh D. Ha, and Eugene V. Koonin. 2021. “A Phylogenomic Framework for Charting the Diversity and Evolution of Giant Viruses.” *PLOS Biology* 19 (10): e3001430. https://doi.org/10.1371/journal.pbio.3001430.

Bhattacharjee, Ananda S., Frederik Schulz, Tanja Woyke, Beth N. Orcutt, and Joaquín Martínez Martínez. 2023. “Genomics Discovery of Giant Fungal Viruses from Subsurface Oceanic Crustal Fluids.” *ISME Communications* 3 (1): 1–9. https://doi.org/10.1038/s43705-022-00210-8.

Cantino, Edward, and Lovett. 1964. “Non-Filamentous Aquatic Fungi: Model Systems for Biochemical Studies of Morphological Differentiation.” In *Advances in Morphogenesis*, 3:33–93. Elsevier. https://doi.org/10.1016/B978-1-4831-9950-4.50005-9.

Clemons, Rebecca, Mark Yacoub, Evelyn Faust, L. Felipe Toledo, Thomas S. Jenkinson, Erik Kalinka, Lillian K. Fritz-Laylin, Timothy Y. James, and Jason E. Stajich. 2023. “DNA Virus BdDV-1 of the Amphibian Pathogen Batrachochytrium Dendrobatidis Is Associated with Hypervirulence.” *bioRxiv*, March, 2023.03.16.532857. https://doi.org/10.1101/2023.03.16.532857.

Cock, J. Mark, Lieven Sterck, Pierre Rouzé, Delphine Scornet, Andrew E. Allen, Grigoris Amoutzias, Veronique Anthouard, et al. 2010. “The Ectocarpus Genome and the Independent Evolution of Multicellularity in Brown Algae.” *Nature* 465 (7298): 617–21. https://doi.org/10.1038/nature09016.

Emerson, R. (1941) An Experimental Study of the Life Cycles and Taxonomy of Allomyces. Lloydia, 4, 77-144.

Frank, JA, and Feschotte. 2017. “Co-Option of Endogenous Viral Sequences for Host Cell Function.” *Current Opinion in Virology* 25 (August): 81–89. https://doi.org/10.1016/j.coviro.2017.07.021.

Gallot-Lavallée, L, and G Blanc. 2017. “A Glimpse of Nucleo-Cytoplasmic Large DNA Virus Biodiversity through the Eukaryotic Genomics Window.” Viruses 9 (1). https://doi.org/10.3390/v9010017.

Gong, Zhen, Yu Zhang, and Guan-Zhu Han. 2020. “Molecular Fossils Reveal Ancient Associations of dsDNA Viruses with Several Phyla of Fungi.” *Virus Evolution* 6 (1). https://doi.org/10.1093/ve/veaa008.

Grigoriev, Igor V., Roman Nikitin, Sajeet Haridas, Alan Kuo, Robin Ohm, Robert Otillar, Robert Riley, et al. 2014. “MycoCosm Portal: Gearing up for 1000 Fungal Genomes.” *Nucleic Acids Research* 42 (D1): D699–704. https://doi.org/10.1093/nar/gkt1183.

Henry, Eric C., and Russel H. Meints. 1992. “A PERSISTENT VIRUS INFECTION IN FELDMANNIA (PHAEOPHYCEAE)1.” *Journal of Phycology* 28 (4): 517–26. https://doi.org/10.1111/j.0022-3646.1992.00517.x.

James, Timothy Y., Jason E. Stajich, Chris Todd Hittinger, and Antonis Rokas. 2020. “Toward a Fully Resolved Fungal Tree of Life.” Review-article. Https://Doi.Org/10.1146/Annurev-Micro-022020-051835. Annual Reviews. World. September 9, 2020. https://doi.org/10.1146/annurev-micro-022020-051835.

Li, Pengfei, Shuangchao Wang, Lihang Zhang, Dewen Qiu, Xueping Zhou, and Lihua Guo. 2020. “A Tripartite ssDNA Mycovirus from a Plant Pathogenic Fungus Is Infectious as Cloned DNA and Purified Virions.” *Science Advances* 6 (14): eaay9634. https://doi.org/10.1126/sciadv.aay9634.

Machlis, Leonard. 1969. “Fertilization-Induced Chemotaxis in the Zygotes of the Watermold Allomyces.” *Physiologia Plantarum* 22 (2): 392–400. https://doi.org/10.1111/j.1399-3054.1969.tb07391.x.

Medina, Edgar M., Jonathan J. Turner, Raluca Gordân, Jan M. Skotheim, and Nicolas E. Buchler. 2016. “Punctuated Evolution and Transitional Hybrid Network in an Ancestral Cell Cycle of Fungi.” eLife. eLife Sciences Publications Limited. May 10, 2016. https://doi.org/10.7554/eLife.09492.

Meints, Russel H., Richard G. Ivey, Amy M. Lee, and Tae-Jin Choi. 2008. “Identification of Two Virus Integration Sites in the Brown Alga Feldmannia Chromosome.” *Journal of Virology*, February. https://doi.org/10.1128/jvi.01983-07.

Moniruzzaman, Mohammad, Carolina A. Martinez-Gutierrez, Alaina R. Weinheimer, and Frank O. Aylward. 2020. “Dynamic Genome Evolution and Complex Virocell Metabolism of Globally-Distributed Giant Viruses.” *Nature Communications* 11 (1): 1–11. https://doi.org/10.1038/s41467-020-15507-2.

Moniruzzaman, Mohammad, Alaina R. Weinheimer, Carolina A. Martinez-Gutierrez, and Frank O. Aylward. 2020. “Widespread Endogenization of Giant Viruses Shapes Genomes of Green Algae.” *Nature* 588 (7836): 141–45. https://doi.org/10.1038/s41586-020-2924-2.

Müller, D. G., and Elisa Parodi. 1993. “Transfer of a Marine DNA Virus fromEctocarpus toFeldmannia (Ectocarpales, Phaeophyceae): Aberrant Symptoms and Restitution of the Host.” *Protoplasma* 175 (3): 121–25. https://doi.org/10.1007/BF01385009.

Myers, J. M., A. E. Bonds, R. A. Clemons, N. A. Thapa, D. R. Simmons, D. Carter-House, J. Ortanez, et al. 2020. “Survey of Early-Diverging Lineages of Fungi Reveals Abundant and Diverse Mycoviruses.” *mBio* 11 (5): e02027-20. https://doi.org/10.1128/mBio.02027-20.

Olsen, Lauritz. 1984. *Allomyces: A Different Fungus*. Opera Botanica 73. Council for Nordic Publications in Botany.

Phadke, Sujal S., Shawn M. Rupp, and Melissa A. Wilson Sayres. 2018. “Evolution of Anisogamy in the Early Diverging Fungus, Allomyces.” *bioRxiv*, November, 230292. https://doi.org/10.1101/230292.

Saito, Makoto, Peiyu Xu, Guilhem Faure, Samantha Maguire, Soumya Kannan, Han Altae-Tran, Sam Vo, AnAn Desimone, Rhiannon K. Macrae, and Feng Zhang. 2023. “Fanzor Is a Eukaryotic Programmable RNA-Guided Endonuclease.” *Nature* 620 (7974): 660–68. https://doi.org/10.1038/s41586-023-06356-2.

Schulz, Frederik, Lauren Alteio, Danielle Goudeau, Elizabeth M. Ryan, Feiqiao B. Yu, Rex R. Malmstrom, Jeffrey Blanchard, and Tanja Woyke. 2018. “Hidden Diversity of Soil Giant Viruses.” *Nature Communications* 9 (1): 1–9. https://doi.org/10.1038/s41467-018-07335-2.

Schulz, Frederik, Simon Roux, David Paez-Espino, Sean Jungbluth, David A. Walsh, Vincent J. Denef, Katherine D. McMahon, et al. 2020. “Giant Virus Diversity and Host Interactions through Global Metagenomics.” *Nature* 578 (7795): 432–36. https://doi.org/10.1038/s41586-020-1957-x.

Starrett, Gj, Mj Tisza, Nl Welch, Ak Belford, A Peretti, Dv Pastrana, and Cb Buck. 2021. “Adintoviruses: A Proposed Animal-Tropic Family of Midsize Eukaryotic Linear dsDNA (MELD) Viruses.” *Virus Evolution* 7 (1). https://doi.org/10.1093/ve/veaa055.

Stevens, Kim, Karen Weynberg, Christopher Bellas, Sonja Brown, Colin Brownlee, Murray T. Brown, and Declan C. Schroeder. 2014. “A Novel Evolutionary Strategy Revealed in the Phaeoviruses.” *PLOS ONE* 9 (1): e86040. https://doi.org/10.1371/journal.pone.0086040.

Swafford, Andrew J. M., and Todd H. Oakley. 2018. “Multimodal Sensorimotor System in Unicellular Zoospores of a Fungus.” *Journal of Experimental Biology* 221 (2). https://doi.org/10.1242/jeb.163196.

Yu, Xiao, Bo Li, Yanping Fu, Daohong Jiang, Said A. Ghabrial, Guoqing Li, Youliang Peng, et al. 2010. “A Geminivirus-Related DNA Mycovirus That Confers Hypovirulence to a Plant Pathogenic Fungus.” *Proceedings of the National Academy of Sciences* 107 (18): 8387–92. https://doi.org/10.1073/pnas.0913535107.

Yutin, Natalya, Yuri I. Wolf, Didier Raoult, and Eugene V. Koonin. 2009. “Eukaryotic Large Nucleo-Cytoplasmic DNA Viruses: Clusters of Orthologous Genes and Reconstruction of Viral Genome Evolution.” *Virology Journal* 6: 223. https://doi.org/10.1186/1743-422X-6-223.

Zhao, Hongda, Ruixuan Zhang, Junyi Wu, Lingjie Meng, Yusuke Okazaki, Hiroyuki Hikida, and Hiroyuki Ogata. 2023. “A 1.5-Mb Continuous Endogenous Viral Region in the Arbuscular Mycorrhizal Fungus Rhizophagus Irregularis.” *Virus Evolution* 9 (2). https://doi.org/10.1093/ve/vead064.

Fungal strains, nucleic acid preparation, and sequencing:

We initially searched 135 genomes in this study including 131 fungi, 92 of which are zoosporic *(Table S1)*. For full details about DNA preparation and sequencing of these isolates see (Amses et al. 2022). An additional nine strains of *Allomyces* sp. were whole-genome sequenced from DNA archived in the CZEUM collection (Simmons et al. 2020). Library preparation and sequencing was performed by the Advanced Genomics Core at the University of Michigan by NovaSeq S4 300 cycle on 4% of a flowcell.

Allomyces javanicus Cali12 and *Allomyces arbusculus* Burma1F were grown in PmTG liquid media for 7-10 days at 22℃ with shaking at 180 rpm. Tissue was subsampled with sterile tweezers and immediately frozen and ground to a powder in liquid nitrogen. Approximately 50 mg was apportioned to 1.5mL tubes, 1 mL of TRIzol applied, and stored at -20℃ until further processing. For *A. arbusculus* Burma1F, the remaining tissue was decanted, rinsed once with sterile water, and replenished with sterile water. After overnight incubation at 22℃, tissue was decanted, rinsed once with sterile water, subsampled and processed as described above, then transferred to a sterile tea strainer and submerged in 50mL sterile water for two hours on the benchtop. The tea strainer was removed, and the remaining liquid observed under a microscope for zoospore presence, followed by centrifugation at 4000x g and 4℃ for 12 minutes. The zoospore pellet was decanted, 1mL of TRIzol added, ground with a pestle, then stored at -20℃. The RNA extraction for all samples was completed per manufacturer protocol. Total RNAs were quality-checked on a Tapestation and had RIN values between 6.8-10. Samples were Poly-A enriched, and libraries were prepared and sequenced by the University of Michigan Advanced Genomics Core by NovaSeq S4 300 cycle on 1% of the flowcell.

Genome and transcriptome assembly, annotation, and statistics:

Transcriptome data were quality filtered by removing reads with Phred score <20 using Fastx-toolkit (Hannon 2010), then assembled *de novo* using Trinity 2.12 (Grabherr et al. 2011). Hisat2 v.2.2.1 (Kim et al. 2019) was used for the RNAseq analysis: the hisat2-build function was used to build an index from genome assemblies, then RNA sequencing reads were aligned to the indexed assemblies using default parameter settings. Alignments were visualized with SeqMonk v.1.48.1 (RRID:SCR_001913).

Paired-end genomic data of *Allomyces* strains were quality-filtered by trimmomatic (Bolger, Lohse, and Usadel 2014) with a sliding window of 4bp and threshold of Phred 20, then assembled by Spades 3.15.5 (Prjibelski et al. 2020) and annotated with both Prodigal v. 2.6.3 (Hyatt et al. 2010) and Augustus 3.2.1 (Stanke et al. 2006) trained with *Rhizopus oryzae*.

To identify likely NCLDV regions or contigs, summary statistics were computed on a per-contig basis. The statistics included depth of read coverage, gene density, gene length, ratio of forward and reverse strand genes, average intergenic distances, fraction of overlapping open reading frames predicted, G+C percentage, intron abundance, number of significant matches to NCVOG and PFAM databases, and number of Capsid, Ribosomal rDNA, and Mitochondrial

genes found, per contig. The scripts and pipeline to compute these statistics are archived in the github repository (https://github.com/stajichlab/Chytrid_Virus) and Stajich 2024.

For the genome plots of *Allomyces arbusculus* Burma 1F *and Allomyces javanicus* California12, we used Prodigal v. 2.6.3 (Hyatt et al. 2010) for gene prediction of soft-masked viral contigs with inputs from Funannotate v. 1.8.16 (Palmer and Stajich 2023) which is a fungal gene annotation and GeneMarkS (http://exon.gatech.edu/genemark/index.html) using the viral option for input sequence type. GC skew was calculated in a 5-kb sliding window with a 1-kb step size. Circos v. 0.54 (Krzywinski et al. 2009) was used to construct the circular genome visualization.

Identification of viral content:

We used the program ViralRecall (Aylward and Moniruzzaman 2021) to identify putative viral genes, viral contigs, and viral regions within contigs. We first ran ViralRecall 1.0 on all assemblies with both stringent (-s 15 -m 30 -v 10) and relaxed (-s 15 -m 2 -v 2) parameters–the latter for poorer quality genome assemblies–and identified isolates containing MCP homologs with e-value <$1e^{-10}$. ViralRecall searches gene predictions generated using Prodigal (Hyatt et al. 2010); we also performed hmmsearches (HMMER3; S. Eddy 2009)with the ViralRecall 2.0 GVOG hmm on a subset of our Augustus-predicted genomes. We found no differences in the number of MCPs called, and so continued characterizations for the isolates that these initial ViralRecall results predicted at least one MCP homolog. For each MCP-containing contig, we calculated GC content, intron density per the Augustus annotation, and performed BLAST searches (Altschul et al. 1990) on both gene prediction datasets.

To aid in distinguishing viral genomes from within the fungal assemblies, we searched the prodigal annotated genomes the 16 genomes previously identified to contain MCPs with HMMER3 (Eddy 2009) as in Schulz et al. 2020: using hmms of a subset of the nucleocytoplasmic viral orthologous genes (NCVOGs)— the 20 NCVOGs most likely to have been vertically inherited (Yutin et al. 2009). We made individual gene trees including all hits with e-values less than 1e-10 by aligning gene sequences with MAFFT version 7 using the E-INS-i algorithm (Katoh and Standley 2013), trimming the resulting alignments with the -automated1 method in TrimAl (Capella-Gutiérrez 2009), and reconstructing trees with the approximately maximum-likelihood approach implemented in FastTree (Price, Dehal, and Arkin 2009) with 100 bootstrap replicates, rooting at the midpoint. We also ran the latest version of ViralRecall (2.0; Aylward and Moniruzzaman 2021) on these 16 assemblies and used these results in combination with genome statistics to discern putatively viral contigs, which we classified as belonging to a viral genome if they met the following criteria: ViralRecall score >2, viral region length (per ViralRecall) >70% of the total contig length, and intron density > 25% different from the host average. One isolate appeared to contain two distinct NCLDVs; we used MetaBAT2 (Kang et al. 2019) on the entire genome of *Allomyces javanicus* California12 under the default settings to separate the two NCLDV genomes.

We conducted BLASTP searches of the hybrid proteomes against nr to identify homologs (e-value <1e-10). Because many symbiont genes are likely misidentified as their hosts' in the

databases, we labeled a gene as “viral” or “bacterial” if any of the top hits were to a virus or bacterium. Some genes have homologs only in the fungi which we have shown also encode MCP; we labeled these “putative FNCLDV”.

To assess prevalence of fungal giant virus-like MCP throughout the fungal kingdom, we used blastp to search all unmasked fungal genomes in Mycocosm (Grigoriev et al. 2014) using MCP of *Allomyces arbusculus* Burma1F as query (total 2,543 genomes).

Giant virus phylogeny, orthology analysis, and KilA-N domain search:
To build the giant virus species tree the nsgtree pipeline was used (https://github.com/NeLLi-team/nsgtree) on genome representatives of the viral phylum Nucleocytoviricota (Aylward et al. 2021). In brief, 7 giant virus orthologous groups (GVOGs; Aylward et al. 2021) were identified using hmmsearch, extracted GVOGs were aligned with MAFFT (version 7.31; Katoh and Standley 2013), trimmed with trimal (-gt 0.1, v1.4; Capella-Gutiérrez, Silla-Martínez, and Gabaldón 2009) and concatenated. To limit the amount of missing data we removed genomes that had less than 4 out of 7 GVOGs. A species tree was built from the supermatrix alignment using IQ-Tree (version 2.03); (Minh et al. 2020; Kalyaanamoorthy et al. 2017) with LG+I+F+G4 and ultrafast bootstrap replications (Hoang et al. 2018) and visualized in iToL (Letunic and Bork 2021).

To analyze the global distribution of *Mycodnaviridae* we used the MCP as a query for a blastp using diamond (--query-cover 30 --subject-cover 30; Buchfink, Xie, and Huson 2015) against all MCP that we previously extracted (Schulz et al. 2020) from publicly available IMG metagenomes (Chen et al. 2023). All blastp hits were then extracted and used as query for another blastp against a database consisting of all MCP extracted from representative giant virus genomes (Aylward et al. 2021) and the initial *Mycodnaviridae* MCPs. We considered as bidirectional best blastp hits (BBH) the environmental MCP that had the best hit based on bitscore to the initial *Mycodnaviridae* MCPs. Geolocation of the BBHs was then used to plot the distribution of MCPs related to *Mycodnaviridae* on a world map projection.

To search for homologs of the viral KilA-N domain in *Mycodnaviridae* genomes, the source data from figure 6 of Medina et al. 2016 (219 protein sequences) was downloaded. Sequences were aligned using mafft (version 7.31; Katoh and Standley 2013), and an hmmprofile made with HMMER hmmbuild (Eddy 2009), which was then used to hmmsearch the fungal giant virus genomes.

Orthofinder 2.5.2 (Emms and Kelly 2019) was used to identify orthologous gene clusters from the prodigal predictions of the most complete viral genomes, determined as the viral genomes containing the greatest number (at least nine) of the NCVOGs20. Visualization of the shared orthologous gene clusters was by adaptation of https://commons.wikimedia.org/wiki/File:Symmetrical_5-set_Venn_diagram.svg. Sequences of orthologous gene clusters were fed into EggNOG-mapper 5.0 (Huerta-Cepas et al. 2019) with default parameters but enabling SMART annotations.

Allomyces capsid PCR and growth experiment:

Primers were designed from an alignment of the MCP sequences of *A. javanicus* California12 and *A. arbusculus* Burma1F (F 5' TACACCAACTTTGCGATGGA3' R 5' GTAGCTCGTTGACCCGACAT3'). DNA was provided by the Collection of Zoosporic Eufungi at the University of Michigan. A standard PCR was performed using GoTaq (Promega) and the products visualized by agarose gel electrophoresis for absence/presence. Burma1F was used as a positive control, and Burma3-35 served as a negative control based on absence of the gene in the sequenced genome.

We chose 6 strains that tested positive for MCP (BEA2, Burma1F, Cali8, Cali12, FijiF1, NorthCaro2), and 6 testing negative (ATCC10983, Burma3-35, CubaS7, DJ02, DJ07, Mex58) to determine phenotypic effects of the gene's presence. Strains were grown in triplicate on Emerson's YpSS media (Emerson 1958). The inoculum for each plate was made by growing strains for 48 hours on fresh YpSS and removing a small circular inoculum by punching discs with the back of a glass pipette. After 72 hrs of growth at 35 C, linear growth was measured for each plate with two perpendicular radii. We also estimated biomass accumulation of the same strains using liquid YpSS. Triplicate 250 ml flasks with 100 ml of media were inoculated with 4 discs cut using the back of glass pipettes. The strains were grown at 35 C for 48 hours with shaking at 120 rpm. The mycelium was harvested using miracloth filtration, transferred to pre-weighed filter papers, dried completely, and dry weight was measured. A final experiment measured thermal maxima of strains by inoculation of YpSS plates and placing them at multiple temperatures: 38, 40, 42, 44, and 46 C. Duplicate plates were prepared from 9 day old YpSS plates using discs excised using glass pipettes. Plates were scored for evidence of growth, where < 2 mm of growth after 1 week was considered negative.

Transmission Electron Microscopy:

Plate cultures were flooded with WC media for 2 hours. Recovered liquid containing discharged zoospores was fixed with equal volume 3% glutaraldehyde in 0.01M Sym-collidine, incubated for two hours at room temperature, and then overnight at 4℃. Zoospores were pelleted and washed thrice with 0.05% Sym-collidine. The pellet was secondarily fixed with 1% osmium tetroxide at room temperature for 30 minutes, followed by two washes each with 0.05% sym-collidine and water. Samples were embedded in 1% noble agar, cut into small blocks (~1mm^3) washed twice with 10% acetone, and incubated in 2% uranyl acetate at 4℃ overnight. Blocks were serially dehydrated with increasing concentrations of 10%, 30%, 50%, 75%, 95% acetone, followed by two washes with 100% acetone. Samples were infiltrated with a 3:1 acetone:Epon-Araldite resin, and incubated for 1 hr, followed by incubation in 1:1 acetone:resin for 1 hr, and again in 1:3 acetone:resin. Samples were then set in 100% resin, vacuum centrifuged overnight, and polymerized at 60℃ for 48 hours. Embedded material was sectioned at 80nm with a diamond knife, sections collected on 200 mesh formvar/carbon nickel thin bar square mesh, stained with lead citrate, and visualized on a Jeol 1400plus transmission electron microscope at the University of Michigan's Department of Molecular, Cell, and Developmental Biology Imaging Core.

References

Altschul, Stephen, Warren Gish, Webb Miller, Eugene Myers, and David Lipman. 1990. “Basic Local Alignment Search Tool.” *Journal of Molecular Biology* 215 (3): 403–10. https://doi.org/10.1016/S0022-2836(05)80360-2.

Amses, Kevin R., D. Rabern Simmons, Joyce E. Longcore, Stephen J. Mondo, Kensuke Seto, Gustavo H. Jerônimo, Anne E. Bonds, et al. 2022. “Diploid-Dominant Life Cycles Characterize the Early Evolution of Fungi.” *Proceedings of the National Academy of Sciences* 119 (36): e2116841119. https://doi.org/10.1073/pnas.2116841119.

Aylward, Frank O., and Mohammad Moniruzzaman. 2021. “ViralRecall—A Flexible Command-Line Tool for the Detection of Giant Virus Signatures in ‘Omic Data.” *Viruses* 13 (2): 150. https://doi.org/10.3390/v13020150.

Aylward, Frank O., Mohammad Moniruzzaman, Anh D. Ha, and Eugene V. Koonin. 2021. “A Phylogenomic Framework for Charting the Diversity and Evolution of Giant Viruses.” *PLOS Biology* 19 (10): e3001430. https://doi.org/10.1371/journal.pbio.3001430.

Bolger, A, M Lohse, and B Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics (Oxford, England)* 30 (15). https://doi.org/10.1093/bioinformatics/btu170.

Buchfink, Benjamin, Chao Xie, and Daniel H. Huson. 2015. “Fast and Sensitive Protein Alignment Using DIAMOND.” *Nature Methods* 12 (1): 59–60. https://doi.org/10.1038/nmeth.3176.

Capella-Gutiérrez, Salvador, José M. Silla-Martínez, and Toni Gabaldón. 2009. “trimAl: A Tool for Automated Alignment Trimming in Large-Scale Phylogenetic Analyses.” *Bioinformatics* 25 (15): 1972. https://doi.org/10.1093/bioinformatics/btp348.

Chen, I.-Min A., Ken Chu, Krishnaveni Palaniappan, Anna Ratner, Jinghua Huang, Marcel Huntemann, Patrick Hajek, et al. 2023. “The IMG/M Data Management and Analysis System v.7: Content Updates and New Features.” *Nucleic Acids Research* 51 (D1): D723–32. https://doi.org/10.1093/nar/gkac976.

Eddy, Sean. 2009. “A New Generation of Homology Search Tools Based on Probabilistic Inference.” *Genome Informatics. International Conference on Genome Informatics* 23 (1). https://pubmed.ncbi.nlm.nih.gov/20180275/.

Emerson, Ralph. 1958. “Mycological Organization.” *Mycologia* 50 (5): 589–621. https://doi.org/10.2307/3756168.

Emms, David M., and Steven Kelly. 2019. “OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics.” *Genome Biology* 20 (1): 1–14. https://doi.org/10.1186/s13059-019-1832-y.

Grabherr, Manfred G., Brian J. Haas, Moran Yassour, Joshua Z. Levin, Dawn A. Thompson, Ido Amit, Xian Adiconis, et al. 2011. “Full-Length Transcriptome Assembly from RNA-Seq Data without a Reference Genome.” *Nature Biotechnology* 29 (7): 644–52. https://doi.org/10.1038/nbt.1883.

Grigoriev, Igor V., Roman Nikitin, Sajeet Haridas, Alan Kuo, Robin Ohm, Robert Otillar, Robert Riley, et al. 2014. “MycoCosm Portal: Gearing up for 1000 Fungal Genomes.” *Nucleic Acids Research* 42 (D1): D699–704. https://doi.org/10.1093/nar/gkt1183.

Hoang, Diep Thi, Olga Chernomor, Arndt von Haeseler, Bui Quang Minh, and Le Sy Vinh. 2018. “UFBoot2: Improving the Ultrafast Bootstrap Approximation.” *Molecular Biology and Evolution* 35 (2): 518–22. https://doi.org/10.1093/molbev/msx281.

Huerta-Cepas, Jaime, Damian Szklarczyk, Davide Heller, Ana Hernández-Plaza, Sofia K. Forslund, Helen Cook, Daniel R. Mende, et al. 2019. “eggNOG 5.0: A Hierarchical, Functionally and Phylogenetically Annotated Orthology Resource Based on 5090 Organisms and 2502 Viruses.” *Nucleic Acids Research* 47 (D1): D309–14. https://doi.org/10.1093/nar/gky1085.

Hyatt, Doug, Gwo-Liang Chen, Philip F. LoCascio, Miriam L. Land, Frank W. Larimer, and Loren J. Hauser. 2010. “Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification.” *BMC Bioinformatics* 11 (1): 1–11. https://doi.org/10.1186/1471-2105-11-119.

Kalyaanamoorthy, Subha, Bui Quang Minh, Thomas K. F. Wong, Arndt von Haeseler, and Lars S. Jermiin. 2017. “ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates.” *Nature Methods* 14 (6): 587–89. https://doi.org/10.1038/nmeth.4285.

Kang, Dongwan D., Feng Li, Edward Kirton, Ashleigh Thomas, Rob Egan, Hong An, and Zhong Wang. 2019. “MetaBAT 2: An Adaptive Binning Algorithm for Robust and Efficient Genome Reconstruction from Metagenome Assemblies.” *PeerJ* 7. https://doi.org/10.7717/peerj.7359.

Katoh, Kazutaka, and Daron M. Standley. 2013. “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.” *Molecular Biology and Evolution* 30 (4): 772. https://doi.org/10.1093/molbev/mst010.

Kim, Daehwan, Joseph M. Paggi, Chanhee Park, Christopher Bennett, and Steven L. Salzberg. 2019. “Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype.” *Nature Biotechnology* 37 (8): 907–15. https://doi.org/10.1038/s41587-019-0201-4.

Krzywinski, Martin, Jacqueline Schein, İnanç Birol, Joseph Connors, Randy Gascoyne, Doug Horsman, Steven J. Jones, and Marco A. Marra. 2009. “Circos: An Information Aesthetic for Comparative Genomics.” *Genome Research* 19 (9): 1639. https://doi.org/10.1101/gr.092759.109.

Letunic, Ivica, and Peer Bork. 2021. “Interactive Tree Of Life (iTOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation.” *Nucleic Acids Research* 49 (W1): W293–96. https://doi.org/10.1093/nar/gkab301.

Medina, Edgar M., Jonathan J. Turner, Raluca Gordân, Jan M. Skotheim, and Nicolas E. Buchler. 2016. “Punctuated Evolution and Transitional Hybrid Network in an Ancestral Cell Cycle of Fungi.” eLife. eLife Sciences Publications Limited. May 10, 2016. https://doi.org/10.7554/eLife.09492.

Minh, Bui Quang, Heiko A. Schmidt, Olga Chernomor, Dominik Schrempf, Michael D. Woodhams, Arndt von Haeseler, and Robert Lanfear. 2020. “IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era.” *Molecular Biology and Evolution* 37 (5): 1530–34. https://doi.org/10.1093/molbev/msaa015.

Palmer, Jonathan, and Jason E. Stajich. 2023. “Funannotate,” August. https://doi.org/10.5281/zenodo.1134477.

Katoh, Kazutaka, and Daron M. Standley. 2013. “MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability.” *Molecular Biology and Evolution* 30 (4): 772. https://doi.org/10.1093/molbev/mst010.

Price, Mn, Ps Dehal, and Ap Arkin. 2009. “FastTree: Computing Large Minimum Evolution Trees with Profiles Instead of a Distance Matrix.” *Molecular Biology and Evolution* 26 (7). https://doi.org/10.1093/molbev/msp077.

Prjibelski, Andrey, Dmitry Antipov, Dmitry Meleshko, Alla Lapidus, and Anton Korobeynikov. 2020. “Using SPAdes De Novo Assembler.” *Current Protocols in Bioinformatics* 70 (1): e102. https://doi.org/10.1002/cpbi.102.

Schulz, Frederik, Simon Roux, David Paez-Espino, Sean Jungbluth, David A. Walsh, Vincent J. Denef, Katherine D. McMahon, et al. 2020. “Giant Virus Diversity and Host Interactions through Global Metagenomics.” *Nature* 578 (7795): 432–36. https://doi.org/10.1038/s41586-020-1957-x.

Simmons, D. Rabern, Anne E. Bonds, Buck T. Castillo, Rebecca A. Clemons, Alex D. Glasco, Jillian M. Myers, Natasha Thapa, et al. 2020. “The Collection of Zoosporic Eufungi at the

University of Michigan (CZEUM): Introducing a New Repository of Barcoded Chytridiomyceta and Blastocladiomycota Cultures." *IMA Fungus* 11 (1): 1–22. https://doi.org/10.1186/s43008-020-00041-z.

Stajich, Jason. Stajichlab/Chytrid_Virus: Code Accompanying Myers et Al (version v1.0). Zenodo, 2024. doi:10.5281/ZENODO.10459876.

Stanke, M, O Keller, I Gundez, A Hayes, Waack, and K Morgenstern. 2006. "AUGUSTUS: Ab Initio Prediction of Alternative Transcripts." *Nucleic Acids Research* 34 (Web Server issue). https://doi.org/10.1093/nar/gkl200.

Yutin, Natalya, Yuri I. Wolf, Didier Raoult, and Eugene V. Koonin. 2009. "Eukaryotic Large Nucleo-Cytoplasmic DNA Viruses: Clusters of Orthologous Genes and Reconstruction of Viral Genome Evolution." *Virology Journal* 6: 223. https://doi.org/10.1186/1743-422X-6-223.