The Diversity of Mycoviruses in Early-Diverging Fungi, and Their Evolutionary Implications

by

Jillian M. Myers

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Ecology and Evolutionary Biology)
in the University of Michigan
2020

Doctoral Committee:

Professor Timothy James, Chair
Assistant Professor Melissa Dude Duhaime
Associate Professor Adam Lauring
Associate Professor Annette Ostling

Jillian M. Myers

jimyers@umich.edu

ORCID iD: 0000-0002-8708-7457

Dedication

For my parents:

By means distinct from–but complementary to–one another,

they each provisioned me with the skills to set and accomplish big goals.

To all others who helped shape my path,

in ways large and small, intentional and serendipitous.

And to all the children in my life.

Acknowledgements

I am most grateful to my advisor, Tim. He is absurdly intelligent, which enables–or perhaps requires–him to take on students whose interests are all over the map, while he continually drives fungal research forward through his own projects. The result is an incredibly intellectually stimulating environment, where one feels surrounded by *big science* most all the time. That could be intimidating. But, as anyone who has ever met Tim knows, he is one of the nicest, most humble humans. He makes a point to build a "mycofamily" of good people. I treasure memories of lab retreats, happy hours, and lab gatherings.

I have not been an easy mentee. I am so grateful to Tim for always reassuring that he believes in me, trusts me, and knows I can do it. That has, literally, made all the difference.

I thank my committee for their support through drastic twists and turns in my research plans. Melissa, Adam, and Annette have each been so willing to help through conversation, networking, and advocacy.

I thank members of the James Lab, past and present. Particularly Alisha Quandt, William Davis, Rabern Simmons, Tommy Jenkinson, Anat Belasen, Kevin Amses, and Rob Powers.

I thank friends in the mycological community for inspiration and inclusion.

There are so many people in EEB I am grateful for that I have decided not to leave the department.

I thank Garret, because without his sacrifices, love, and support this would have been a much darker experience, were it possible at all. And Therrien, for his patience. We can play now.

Table of Contents

List of Tables

List of Figures

Abstract

Even microbes can, themselves, be host to microorganisms. Bacterial-fungal interactions have increasingly received attention, empowered by the identification of widespread endobiotic bacteria in some early-diverging fungal lineages. But, rarely are fungi thought of as hosts to viruses. In this dissertation research, I explore methods for fungal virus (mycovirus) discovery, identify many new viruses, and use these newly obtained viral sequences to address outstanding questions in the field. Specifically, I ask "Which fungi are infected with viruses?", "What kinds of viruses?", and "Where did fungal viruses come from?". I speculate on the implications of viral infection to fungal evolution, fungal ecology, and ecosystems.

One major limitation in mycovirus research, and the focal point of this dissertation, is the lack of diverse fungal sampling. In Chapter 2, through a survey of both cultures and transcriptomes, I find higher viral prevalences in the basal lineages than have previously been reported. I identify at least 85 previously unknown viruses that span RNA virus taxonomy, and demonstrate that mycoviruses are present in research cultures around the world, the implications of which are far-reaching.

These newly identified mycovirus sequences from basal fungi enable a previously impossible examination of the evolutionary relationships between mycoviruses and fungi. In Chapter 3, I test the hypothesis of cospeciation of mycoviruses and their fungal hosts through statistical tests of phylogenetic congruence. I find evidence of cospeciation in all four viral families tested, despite this mode of speciation being quite rare. When evidence of cospeciation does exist, it is most often between hosts and their vertically-transmitted mutualists. There is

evidence that host-switching is the dominant mode of speciation among RNA viruses, and so my findings suggest that mycovirus life history is unique, and of ecological importance.

Known mycoviruses are almost exclusively composed of RNA genomes. In Chapter 4, the number of known mycoviruses with DNA genomes is more than quadrupled with the identification, for the first time in Fungi, of nucleocytoplasmic large DNA viruses (NCLDV). This exciting finding has unknown but important implications to global nutrient cycling as NCLDVs notoriously reprogram their host's metabolism. Importantly, these large viruses likely only occur in some of the most basal, zoosporic, lineages.

As they did for bacterial-fungal interactions, the early-diverging fungi have transformed what we think we know about fungal viruses, demonstrating, once again, that representation shapes knowledge. This work has addressed some of the outstanding questions in the field, and, in the process, raised many more.

Chapter 1 Introduction

The scope of fungal diversity has been underappreciated

Fungi have been celebrated for thousands of years. Through their use as tinder and to transport fire, for nutrition and medicine, fungi have been critical to human survival (Berihuete-Azorín 2018, Vaidya & Rabba 1993, Money 2018a). Yeast-fermented foods and beverages have been crucial parts of human diets over the entire world, independently arising in the form of *chicha* by peoples of the Andes and Amazonia regions of South America, *keribo* in Ethiopia, millet beer of the Senufo, the Lobi, and Koulango communities in West Africa (known as *tchapalo* in Cote d'Ivorie), *koji* products in Japan such as miso paste, sake, and tamari, and the fermented banksia flowers by people indigenous to Australia. Arguably, fungi and humans are inextricably linked, the manifestations of which are abundant in cultural traditions worldwide. The most popularly celebrated fungi, however, are just two types: mushrooms and yeasts. Mushrooms are delicious, richly diverse, and culturally important. Yeasts, arguably, are the most economically important microorganisms (Money 2018b). But the fungal kingdom has even more to offer, though the other groups have received relatively little attention.

Traditional knowledge has included the use of multifarious fungal organisms for biological control, medicines, dyes, and food fermentation, but Western European science has largely neglected other fungal lineages. Mycological studies, even, have underappreciated the inimitable "water molds", rich with complex structure and life cycles for "primitive" fungi, and the duplicity of the "pin molds", which both rot our fruits and transform soybeans into palatable

and nutritious tempeh. The onus has been on a mere handful of scientists whose genuine passion for these organisms has been the single driving force keeping their study current and forward moving. Fortunately, this passion has been generationally inherited. James Lovett describes this, in a *memorium* for his PhD advisor, EC Cantino, as being "lured by the siren's song of water molds", a process mediated through the enthusiasm of one's mentor (Lovett 1984). It's an endearing remembrance, if not relatable to all readers of this dissertation.

These passionate individuals tirelessly developed the field and, as it is with basic science, a time came that necessitated the foundational knowledge that had been established. In the mid 1990s, scientists began noticing entire populations of frogs mysteriously dying, what's more, in seemingly pristine natural environments (Berger 1998). A highly unusual organism was found infecting the skin tissue of the dead frogs. Joyce Longcore, who had fallen prey to the allure of unusual fungi during her own tenure at the University of Michigan, identified this organism as a chytrid fungus shortly after its discovery (Longcore 1999). It is disquieting to imagine an alternative history devoid of the expertise and foundational research that enabled the timely identification and laboratory manipulation of the causative agent of what was one of the most catastrophic wildlife diseases in recorded history. It is our great pleasure and privilege, as "basic" scientists, to study aspects of the natural world buried in obscurity. We do so because, without a doubt, there will be a time where this knowledge, built from such seemingly inconsequential work, will be influential.

Fungi are microbes *and* hosts

Thanks, in no small part, to the amphibian-killing chytrid fungus, the lesser celebrated fungal lineages have begun to be more widely appreciated for their magnificent diversity and consequential ecological interactions. One burgeoning field of research is the study of bacterial-

fungal interactions, which has been propelled forth by the discovery of endobiotic bacteria harbored by multiple groups of these lesser-known fungi. A rhyme by mathematician Augustus de Morgan, written in 1872, begins “Great fleas have little fleas upon their backs to bite ‘em/ And little fleas have lesser fleas, and so *ad infinitum*.” (de Morgan 1972). As he playfully predicts, though most fungi are microbes themselves, they, too, can be host to microorganisms. Recent research on endobiotic bacteria in the pin molds has revealed their widespread distribution, complex interactions, and common patterns of bacterial genome reduction (Bonfante and Desirò 2017).

Much has been learned by appreciating that these fungi are host to bacteria, but they have not yet been explored as hosts of viruses. The research presented herein is an expedition in discovering and describing viruses in these underappreciated fungal lineages and characterizing their historical relationships through sequence-based evolution. Fundamentally, this dissertation demonstrates that these fungal lineages have even more to teach us about fungal-microbial interactions through transforming what we think we know about fungal viruses.

Viruses

Viruses are wildly abundant and diverse

The definition of “virus” is not concrete and has often required revision once a new virus is discovered that breaks the bounds of the previous definition. For the purposes of this text, we will consider a virus to be a noncellular infectious agent with DNA or RNA genetic material that can direct its own reproduction and spread by co-opting the machinery of the host cell it inhabits.

Viruses are, arguably, the most abundant and diverse genetic entities on the planet (Koonin 2010). This is perhaps most clear by considering the variety in type and arrangement of nucleic acids. It is obvious that there is great genetic distance between humans and amoebae,

measurable as similarity of DNA sequence. Viruses take this to the next level: variation is such that DNA sequences cannot even be compared, because homologous sequences, *if* they exist, can be so dissimilar as to be unalignable. Viral genetic material can consist of single-stranded DNA (ssDNA), double-stranded DNA (dsDNA), linear positive-sense single-stranded RNA (ssRNA(+); positive-sense refers to RNA polarity such that ssRNA(+) can function as mRNA), linear negative-sense single-stranded RNA (ssRNA(-); negative-sense RNA is complementary to mRNA and must be converted prior to replication), double-stranded RNA (dsRNA), or circular ssRNA. Genome sizes range from the tiny 1.8kb circoviruses to the largest known *Pandoravirus salinus* at 2.5MB, a one-thousand-fold difference. Evolution has produced an impressive array of morphological variation, as well.

Virus-host interactions span the parasitism/mutualism continuum and have ecosystem-level consequences

Interactions between viruses and their hosts are also highly variable. Viruses have conventionally been considered, and at one point defined, as "obligate parasites", which emphasizes their role in disease. While, indeed, viruses rely on their hosts for replication and thus inherently incur a cost, this does not exempt viruses from potentially having a net benefit to their hosts. Indeed, even mutualisms incur a cost (Bronstein 2001). That viruses may occupy the full expanse of the parasitism-mutualism continuum is an ongoing paradigm shift, predominantly sparked by plant and fungal virologists (Roossinck 2005; Márquez and Roossinck 2012; Roossinck 2015; Roossinck 2011). Notably, plant and fungal viruses are often cryptic, persistent, and are frequently vertically inherited, perhaps the signatures of context-dependent mutualisms. A few textbook examples of beneficial viruses have become well-known (Márquez 2007; Schmitt and Breinig 2002).

The effects of viruses are not contained at the level of the individual, but, importantly, also scale up to the ecosystem level. Viral ecology has mostly been driven by marine microbial virologists, particularly through the "viral shunt" and "viral shuttle" paradigms, which describe the mechanisms by which viruses may regulate nutrient fluxes in the oceans (Brussaard 2008; Weitz and Wilhelm 2012; Roux 2016). The consequences of the "viral shunt" are that, through viral lysis of host cells, organic matter is kept small, which reduces the flux of carbon to larger eukaryotes via food as well as from the ocean surface (Wilhelm and Suttle 1999). The "viral shuttle" involves aggregation of viral lysates, which fall and increase carbon flux to the deep ocean (Weinbauer 2004). Indeed, Tara Oceans studies have demonstrated, through genomic surveys and models, that viruses are the best biological predictors of global ocean carbon fluxes (Guidi 2016). Such studies have focused on viruses of microbes, which are abundant in soils, but similar studies at this scale remain to be conducted in terrestrial habitats. Further, the microbial viruses measured and modeled typically only represent those infecting bacteria and archaea, with viruses of protists and fungi largely ignored.

Mycoviruses

Mycoviruses are persistent in their hosts

Viruses that infect fungi are termed mycoviruses. The overwhelming majority of known mycoviruses have RNA genomes and only two DNA viruses have been identified in the entire kingdom, both small, circular replication-protein encoding single stranded (CRESS) viruses (Yu 2010; Li 2020). Approximately two-thirds of known RNA mycoviruses have a dsRNA genome and the majority of the remaining have ssRNA(+) genomes, with a handful of viruses with ssRNA(-) genomes (Ghabrial 2015). Mycoviruses have primarily been reported from plant

pathogens in the Ascomycota, but also in the Basidiomycota and, more recently, other fungal lineages (Papp 2001; Kartali 2019; Coyle 2018; Neupane 2018; Kitahara 2014).

Mycoviruses perhaps frequently go unreported because of their curious "persistent" and "latent" infection strategies. They are often described as persistent because they apparently lack an extracellular route of infection and instead transmit vertically through sexual and asexual spores and intracellularly via cell division and hyphal fusion between compatible individuals. Fungal viruses can be difficult to cure, and can persist in cultures. Unlike plant viruses, many of which are also persistent, they do not encode movement proteins, which modify plant plasmodesmata to allow viral movement between cells. Mycoviral infection is also frequently described as latent because it is usually asymptomatic, though there are many exceptions that include phenotypes such as hypovirulence, hypervirulence, pigmentation changes, and irregular growth. Mycoviruses that induce obvious hypovirulent phenotypes are uncommon but have arguably received the most attention for their potential for biocontrol of fungal pathogens (García-Pedrajas 2019; Pearson 2009). Most notably, the hypoviruses of *Cryphonectria parasitica*, the fungal pathogen responsible for the annihilation of chestnut trees in North America, effectively controlled this pathogen in European trees (Rigling and Prospero 2018). These hypoviruses have been developed into model systems, enabling research on the intricacies of the arms-race between virus and host (Segers 2007; Sun 2009; Shahi 2019; Zhang 2012). However, the hypoviruses are quite exceptional. More frequently, the impacts of viral infection to the host are difficult to assess in the laboratory since phenotypic results of infection are subtle. In nature, whether a mycovirus is asymptomatic, detrimental, or beneficial, may depend on environmental and ecological factors (Hyder 2013). The asymptomatic nature of most mycoviral

infections might suggest a long history of association and, perhaps, conferred benefits, if contextually dependent.

Mycovirus-host interactions span ecological scales and continuums

One particularly unique phenotype is that of the "killer" viruses in yeasts. These vertically inherited dsRNA viruses encode a toxin gene that targets the fungal cell wall, and also encode host immunity to that toxin (Schmitt and Breinig 2002). Thus, under conditions of competition, killer viruses provide a significant competitive advantage to their hosts, and their offspring. The ecological and evolutionary relevance of this competitive advantage was elucidated by Drinnenberg *et al.*, who found, remarkably, that RNA interference (RNAi), a fungal immune defense targeting dsRNA, was independently lost in species host to killer viruses (Drinnenberg 2011). The long-term consequences of RNAi loss remain to be seen but, since this is also the primary defense against selfish transposable elements, it could lead to extinction of these species in the evolutionary long-term (Drinnenberg 2011).

Fungal immune defense to mycoviral infection has thus far has proven quite complex. Since RNAi specifically targets dsRNA, it has been the main interest of researchers interested in host-virus interactions at the molecular scale. Interestingly, there appears to be substantial variability in whether, and the extent to which, RNAi is induced by the presence of mycoviruses. RNAi has definitively been shown to be involved in antiviral defense in *C. parasitica* against the most debilitating hypovirus (Segers et al. 2007; Sun, Choi, and Nuss 2009) but other, more mild, hypovirus strains do not significantly induce expression of a major RNAi player, the *dcl*2 protein(Zhang, Shi, and Nuss 2012). Nor does another lineage of virus, mitoviruses, increase expression of RNAi genes in *C. parasitica* (Shahi 2019). It is also possible that some mycoviruses have evolved counter-defenses in the form of virus-derived small interfering RNAs

(vsiRNAs) that could modulate fungal RNAi defenses and effectively silence them (Özkan 2017; Wang 2016). Overall, it appears that RNAi responses, and potentially viral counter-responses, likely vary according to fungal and/or mycovirus species.

Likewise, overall transcriptional response to infection is highly dependent on the particular mycovirus and host. In *C. parasitica*, hypovirus infection downregulates the expression of transcription factors for proteins involved in the MAPK pathway which are critical for invasive growth and mating (Allen 2003; Allen and Nuss 2004; Deng 2007). Transcriptomics in the plant pathogen *Fusarium graminearum* infected with one of four mycoviruses demonstrated the surprising result that noticeable phenotypic alteration as a result of mycovirus infection did not correlate with impact to host transcriptional profiles (Lee 2014). They found that viruses inducing either asymptomatic or hypovirulent phenotypes could have substantial host transcriptional impact, and also that each of the four mycoviruses tested regulated a different suite of fungal genes. Interestingly, a *Heterobasidion parviporum* mycovirus that causes hypovirulence in this plant pathogen impacts metabolism in the host (Vainio 2017). Specifically, carbohydrate metabolism was affected and alternative metabolic pathways were upregulated. It is compelling to consider the accumulated impacts of such metabolic changes at the ecosystem level, but so far estimates of the impact of fungal viruses to nutrient cycling remains one of the many unanswered questions.

Mycoviral origins are unknown

Another unanswered, but fundamental, question regards the origin of mycoviruses. Two hypotheses, not mutually exclusive of one another, have been suggested (Pearson2009, Son 2015). The “plant virus hypothesis” suggests that virus infecting a plant host-jumped to a fungus on the same plant. The “ancient coevolution hypothesis” suggests that mycoviruses and their

hosts have a long history of association that includes cospeciation. Largely due to lack of diverse mycoviral sampling, neither hypothesis has convincing evidence to support it.

Mycoviruses are understudied but gaining interest

The first mycovirus was identified about only 60 years ago, and much has been learned since (Ghabrial 2015). However, many fundamental aspects remain unanswered. Mycovirus research has not yet come of age (Figure 1.1). Particularly, the majority of mycoviruses identified are in fungi in the two most celebrated fungal groups previously mentioned, and their close relatives (the Dikarya). Only a handful of mycoviruses have been described from the other groups, which comprise the phylogenetic diversity of the kingdom (Papp 2001; Coyle 2018; Kitahara 2014; Marzano 2016). What might be learned by searching for viruses in the more obscure lineages?

Early-diverging lineages of fungi

Kingdom Fungi is made up of at least 8 phyla: Basidiomycota, Ascomycota, Mucoromycota, Zoopagomycota, Blastocladiomycota, Chytridiomycota, Neocallimastigomycota, and Cryptomycota/Rozellomycota (Berbee 2017; Spatafora 2016; Hibbett 2007) (Figure 1.2). Basidiomycota and Ascomycota, together called “Dikarya”, are the two most recently diverged groups and make up roughly 98% of the approximately 144,000 described fungal species (Moore 2011; State of the world’s fungi 2018), though at least 2.2 million species exist (Hawksworth and Lücking 2017). This disparity in description between groups is striking.

The fungal kingdom is remarkably diverse, and synapomorphies are few. Amongst the non-Dikaryotic lineages the only shared characteristic is the lack of regular septa in mycelia (with the exception of the Kickxellomycotina, phylum Zoopagomycota). The following sections

aim to provide a brief overview of each phylum of the non-Dikarya, referred to as the "early-diverging lineages" throughout this text.

Cryptomycota/Rozellomycota

Cryptomycota forms the most basal branch of the fungal tree. They are primarily known from environmental sequence data, where they are abundant (Jones 2011; Lazarus and James 2015). Little is known of their cell biology and ecology as a group, but the organisms that have been directly observed are intracellular parasites of eukaryotes. These include *Rozella* spp., which infect another early-diverging fungus, *Allomyces*, as well as algae and oomycetes, *Nucleophaga* spp., which parasitize amoebae, and *Paramicrosporidium* spp., specializing on the nuclei of amoebae (Corsaro 2014a; Corsaro 2014b; Held 1980). They reproduce via small, sometimes motile zoospores, and lack a cell wall during the trophic phase. A particularly large radiation, this single phylum is perhaps as diverse as the entire fungal kingdom (Jones 2011).

Chytridiomycota

The "chytrids" have a global distribution, with "apparently unlimited modes of survival" (Powell 2016). They are common in aquatic environments (both freshwater and marine), and in soils of all types, including forests, agricultural, deserts, grasslands, acidic bogs, and alpine mountains under snow (Powell 2016; Moore 2011). They are morphologically relatively simple, consisting of a thallus, or the main body of the organism, and rhizoids, or branching elements that anchor the thallus to the substrate. Despite the simplicity, there exist a variety of thallus forms that are adaptive for the various substrates they specialize in or on. Chytrid reproduction involves the thallus converting into a zoosporangium, which divides to produce numerous single-celled zoospores, which lack a cell wall and are motile by means of a posterior flagellum. After a motile period, the flagella retracts, a wall is produced, and a thallus develops.

Chytrids often specialize on substrates that are indigestible by other organisms, such as pollen grains, keratin, chitin, and cellulose. They therefore play a vital role in nutrient cycling. Often saprotrophic, some are also notable parasites of algae. Hassett *et al.* demonstrate that chytrids dominate the fungal community in the arctic marine environment and, mediated by climate change, have the potential to drastically change patterns of primary production through their prey on algae with unknown consequences to reliant trophic cascades (Hassett and Gradinger 2016).

Neocallimastigomycota

These unique fungi are obligate anaerobes living in the guts of ruminants and herbivorous reptiles, where they break down food substrates that would otherwise be indigestible to the host. Thus, they have been "crucial to the evolution of herbivores and the prosperity of animal husbandry since humans domesticated animals" (Moore 2011) and remain ecologically and economically important.

They are chytrid-like morphologically: either singly or multiply flagellated zoospores encyst on plant material in the gut, develop a thallus, and produce degrading enzymes that decompose the substrate, making essential nutrients bioavailable for the host. Unsurprisingly, their unique ecology makes them logistically challenging to study, and so the diversity of the Neocallimastigomycota is likely much greater than is currently characterized.

Blastocladiomycota

The "blastoclads" include saprotrophs, plant parasites, and obligate endoparasites of insects. These, too, are among the zoosporic fungi and have substantial morphological diversity. The most striking feature of this group is their unique and complex life cycle that includes sporic meiosis ultimately resulting in the alternation of generations between haploid and diploid

individuals. Unique and impressive innovations of this “primitive” group include anisogamy and sexual signaling: small “male” gametes are attracted by the pheromones released by large “female” gametes. There is something undeniably alluring about this group, that has sparked the interest of even artists (Figure 1.3).

Zoopagomycota

The Zoopagomycota represent the earliest transition from the ancestral zoosporic fungi previously mentioned to aerially-dispersing, filamentous fungi. They are primarily animal-associates, including parasites and endosymbionts, but include some mycoparasites as well. Especially interesting, this group contains multiple predators with highly specialized structures and strategies for hunting their nematode or amoebae prey. They also include gut endosymbionts of arthropods that are so widespread that “just about every arthropod that crawls past you will carry these fungi within” (Moore 2011).

Mucoromycota

This group is composed of an array of plant-associates: mycorrhizal fungi, root endophytes, plant pathogens, as well as decomposers. Perhaps most iconic are the obligate plant mutualists in Glomeromycotina, also called the arbuscular mycorrhizal fungi. The ecological importance of these fungi cannot be overstated, as they form mutualistic associations with about 80% of land plants, providing these plants with essential minerals efficiently obtained by the mycelium extending outside the plant root. These fungi form highly branched structures, called arbuscules, within root cells themselves. The two other subphyla, Mucoromycotina and Mortierellomycotina, are common rhizospheric soil fungi. These are typically fast-growing and early colonizers, including some household fruit and bread molds. Some are notable for causing disease in humans and other animals. Others have significant industrial application. Curiously,

one common feature of the Mucoromycota is the presence of *Burkholderia*-related endobacteria, which are apparently of ancient association (Bonfante and Venice 2020; Mondo 2012).

Mycoviruses in early-diverging fungi: questions, answers, and more questions

The research presented in this dissertation was sown from the fields of two poorly studied subdisciplines, the fertile substrate for novel revelations. In the following chapters, I explore methods for mycoviral discovery in the early-diverging fungi, identify many new viruses, and use these newly found viruses to address outstanding questions in the field. In doing so, I uncover a wealth of novel diversity and complex interactions that call out for further inquiry.

One prescient issue limiting mycovirus research, and the focal point of this dissertation, is the lack of diverse fungal sampling. At the onset, it appeared that antiquated methods of mycoviral discovery could be one contributing factor. In Chapter 2, I address these two issues by thoroughly screening early-diverging fungal lineages for mycovirus using two approaches: a traditional culture-based method and a more modern transcriptome-mining approach. The culture-based method performed, perhaps surprisingly, well compared to transcriptome-mining. Conservatively, I identified 85 mycoviruses previously unknown, comprising taxa spanning Riboviria (RNA viruses), and demonstrate that mycoviruses are present in laboratory cultures around the world, the implications of which are unknown.

These newly obtained mycovirus sequences from basal fungi empowered tests of evolutionary relationships previously unable to be addressed. In Chapter 3, I test the hypothesis of cospeciation of mycoviruses and their fungal hosts through statistical tests of phylogenetic congruence. The four mycoviral families tested demonstrate evidence of cospeciation, despite this mode of speciation being quite rare, suggesting that mycovirus life history is unique.

In Chapter 4, the early-diverging fungi reveal some of their best kept secrets and the number of DNA viruses known from the fungal kingdom is more than quadrupled. Through searches of fungal genomes, we identify large DNA viruses for the first time in Fungi. This exciting finding has unknown, but possibly important, implications to global nutrient cycling as large DNA viruses notoriously reprogram their host's metabolism. Importantly, these large viruses likely only occur in the most basal, zoosporic, lineages.

Although mycovirological research is gaining momentum, it is still wildly underdeveloped. This work has addressed some of the outstanding questions in the field and raised many more. My great hope for this body of work revealing mycoviruses throughout the fungal kingdom, and the transformative revelations therein, is to contribute to the understanding of fungal ecology and evolution, and in doing so lure mycologists to the study of these fascinating obscurities.

Figures

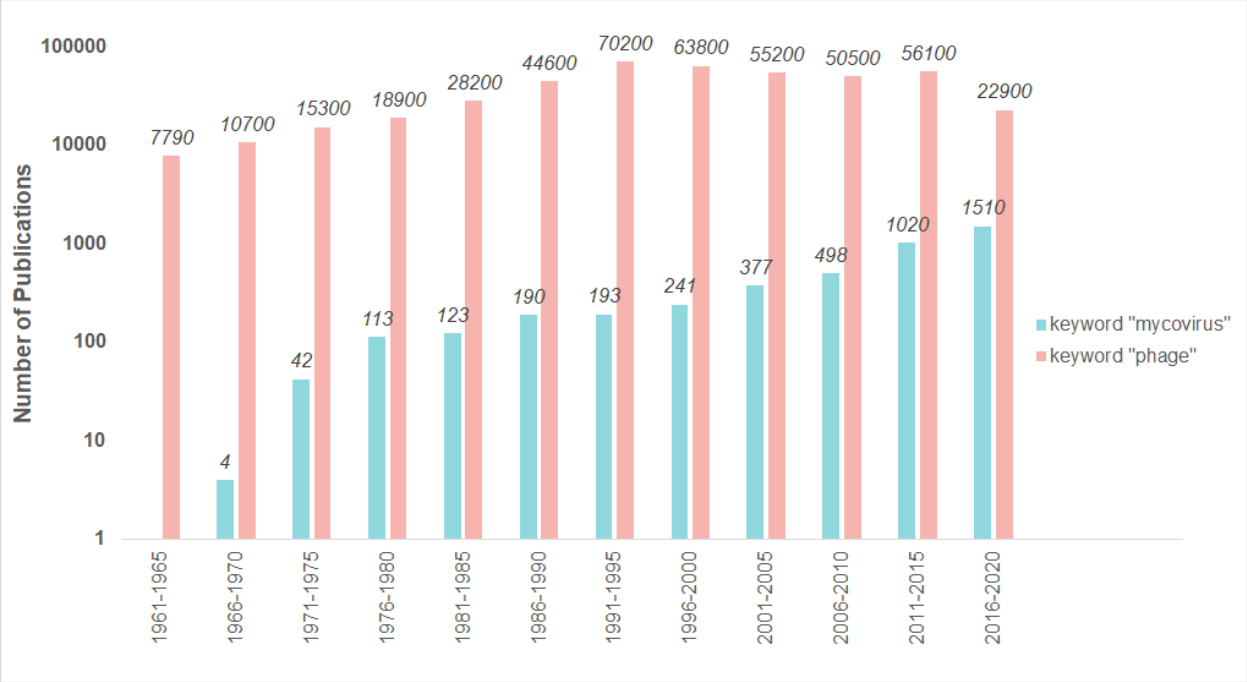


Figure 1.1 Google scholar search using keywords “mycovirus” (teal) or “phage” (pink), binned in 5-year increments. Though the term “phage” refers to viruses infected bacteria, an entire *domain* of life, whereas mycoviruses are found in the fungal *kingdom*, the disparity in number of published papers on the two topics is staggering, apparent by the necessity for a logarithmic scale. However, mycovirus research is experiencing an upward trend in popularity.

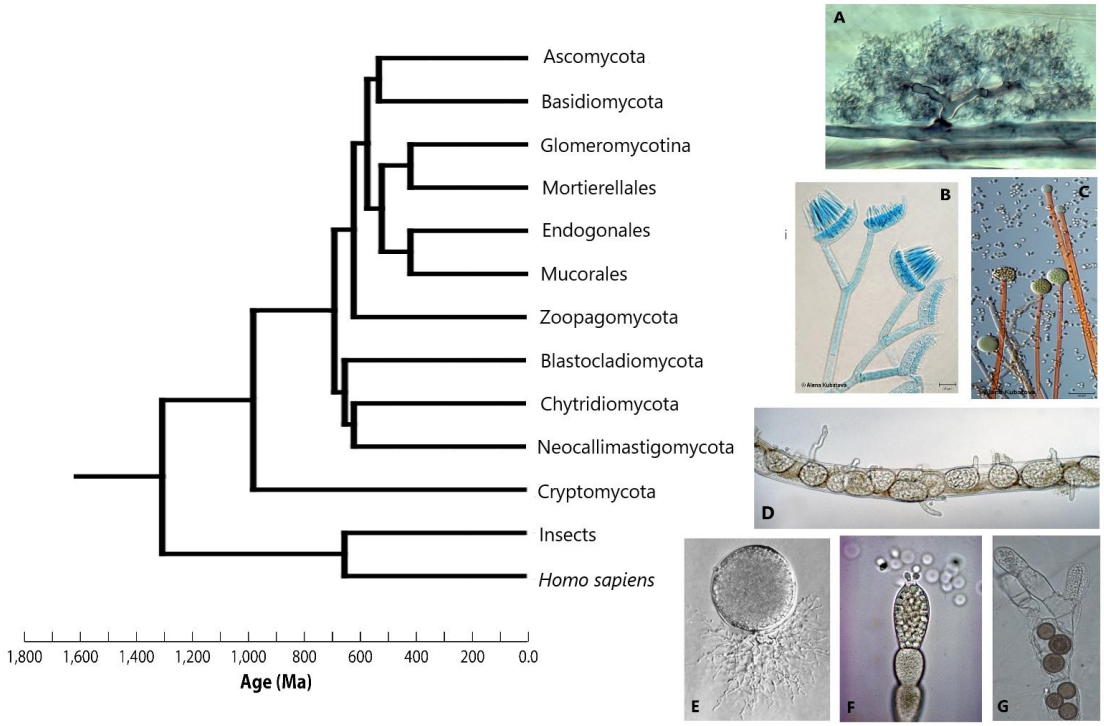


Figure 1.2 Phylogeny of the fungal kingdom, adapted from Berbee 2017. Fungi and Holozoa diverged between 1,481 and 900 million years ago. On left, images of some fungi in early-diverging fungal clades: A) mature arbuscule of *Glomus* within a plant root cell (Glomeromycotina); image copyright Mark Brundrett, 2008 B) *Coemansia reversa* (Zoopagomycota); image copyright Alena Kubátová C) *Umbelopsis ramanianna* (Mucorales); image copyright Alena Kubátová D) *Catenaria anguilullae* (Blastocladiomycota) parasitizing a nematode; image copyright George Barron, 2013 E) *Spizellomyces palustris* (Chytridiomycota); image from Chen 2000 F) *Allomyces* sp. (Blastocladiomycota) sporangia releasing zoospores G) resting spore of *Rozella allomycis* (Cryptomycota) parasitizing *Allomyces* sp.; image by TY James.

Figure 1.3 A painting inspired by Allomyces arbuscula. Watercolor by Debra Myers, 2019.

References

Allen, TD, and DL Nuss. 2004. Specific and common alterations in host gene transcript accumulation following infection of the chestnut blight fungus by mild and severe hypoviruses. *Journal of Virology* **78** (8): 4145–55. https://doi.org/10.1128/JVI.78.8.4145.

Allen, TD, AL Dawe, and DL Nuss. 2003. Use of cDNA microarrays to monitor transcriptional responses of the chestnut blight fungus *Cryphonectria parasitica* to infection by virulence-attenuating hypoviruses. *Eukaryotic Cell* **2** (6): 1253–65. https://doi.org/10.1128/EC.2.6.1253.

Berbee, ML, TY James, and C Strullu-Derrien. 2017. Early diverging fungi: diversity and impact at the dawn of terrestrial life. *Annual Review of Microbiology* **71** (1): 41–60. https://doi.org/10.1146/annurev-micro-030117-020324.

Berger, L, R Speare, P Daszak, DE Green, AA Cunningham, CL Goggin, R Slocombe, *et al*. 1998. Chytridiomycosis causes amphibian mortality associated with population declines in the rain forests of Australia and Central America. *Proceedings of the National Academy of Sciences of the United States of America* **95** (15): 9031–36. https://doi.org/10.1073/pnas.95.15.9031.

Berihuete-Azorín, M, J Girbal, R Piqué, A Palomo, and X Terradas. 2018. Punk's not dead. fungi for tinder at the neolithic site of La Draga (NE Iberia). *PLoS ONE* **13** (4): 1–16. https://doi.org/10.1371/journal.pone.0195846.

Bonfante, P, and A Desirò. 2017. Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. *ISME Journal* **11** (8): 1727–35. https://doi.org/10.1038/ismej.2017.21.

Bonfante, P, and F Venice. 2020. Mucoromycota: going to the roots of plant-interacting fungi. *Fungal Biology Reviews* **34** (2): 100–113. https://doi.org/10.1016/j.fbr.2019.12.003.

Bronstein, JL. 2001. The costs of mutualism. *American Zoologist* **41** (4): 825–39. https://doi.org/10.1093/icb/41.4.825.

Brussaard, CPD, SW Wilhelm, F Thingstad, MG Weinbauer, G Bratbak, M Heldal, SA Kimmance, *et al*. 2008. Global-scale processes with a nanoscale drive: the role of marine viruses. *ISME Journal* **2** (6): 575–78. https://doi.org/10.1038/ismej.2008.31.

Corsaro, D, J Walochnik, D Venditti, KD Müller, B Hauröder, and R Michel. 2014a. Rediscovery of nucleophaga amoebae, a novel member of the Rozellomycota. *Parasitology Research* **113** (12): 4491–98. https://doi.org/10.1007/s00436-014-4138-8.

Corsaro, D, J Walochnik, D Venditti, J Steinmann, KD Müller, and R Michel. 2014b. Microsporidia-like parasites of amoebae belong to the early fungal lineage Rozellomycota. *Parasitology Research* **113** (5): 1909–18. https://doi.org/10.1007/s00436-014-3838-4.

Coyle, MC, CN Elya, M Bronski, and MB Eisen. 2018. Entomophthovirus: an insect-derived iflavirus that infects a behavior manipulating fungal pathogen of dipterans. *BioRxiv*, 371526. https://doi.org/10.1101/371526.

de Morgan A. 1872. A budget of paradoxes. Dover Publications, Inc. New York.

Deng, F, TD Allen, and DL Nuss. 2007. Ste12 transcription factor homologue CpST12 is down-regulated by hypovirus infection and required for virulence and female fertility of the chestnut blight fungus *Cryphonectria parasitica*. *Eukaryotic Cell* **6** (2): 235–44. https://doi.org/10.1128/EC.00302-06.

Drinnenberg, IA, GR Fink, and DP Bartel. 2011. Compatibility with killer explains the rise of RNAi-deficient fungi. *Science* **333** (6049): 1592. https://doi.org/10.1126/science.1209575.

García-Pedrajas, MD, MC Cañizares, JL Sarmiento-Villamil, AG Jacquat, and JS Dambolena. 2019. Mycoviruses in biological control: from basic research to field implementation. *Phytopathology* **109** (11): 1828–39. https://doi.org/10.1094/PHYTO-05-19-0166-RVW.

Lovett, J. 1984. *In Memoriam* Professor Edward C. Cantino. *Experimental Mycology* **8**, 95-104.

Ghabrial, SA, JR Castón, D Jiang, ML Nibert, and N Suzuki. 2015. 50-Plus years of fungal viruses. *Virology* 479–480: 356–68. https://doi.org/10.1016/j.virol.2015.02.034.

Guidi, L, S Chaffron, L Bittner, D Eveillard, A Larhlimi, S Roux, Y Darzi, *et al*. 2016. Plankton networks driving carbon export in the oligotrophic ocean. *Nature* **532** (7600): 465–70. https://doi.org/10.1038/nature16942.

Hassett, BT, and R Gradinger. 2016. Chytrids dominate arctic marine fungal communities. *Environmental Microbiology* **18** (6): 2001–9. https://doi.org/10.1111/1462-2920.13216.

Hawksworth, DL, and R Lücking. 2017. Fungal diversity revisited: 2.2 to 3.8 Million Species. *Microbiology Spectrum* **5**(4) https://doi.org/10.1128/microbiolspec.funk-0052-2016.

Held, AA. 1980. Development of *Rozella* in *Allomyces*: a single zoospore produces numerous zoosporangia and resistant sporangia. *Canadian Journal of Botany* **58** (8): 959–79. https://doi.org/10.1139/b80-119.

Hibbett, DS, M Binder, JF Bischoff, M Blackwell, PF Cannon, OE Eriksson, S Huhndorf, *et al*. 2007. A higher-level phylogenetic classification of the fungi. *Mycological Research*. https://doi.org/10.1016/j.mycres.2007.03.004.

Hyder, R, T Pennanen, L Hamberg, EJ Vainio, T Piri, and J Hantula. 2013. Two viruses of *Heterobasidion* confer beneficial, cryptic or detrimental effects to their hosts in different situations. *Fungal Ecology* **6** (5): 387–96. https://doi.org/10.1016/j.funeco.2013.05.005.

Jones, MDM, TA Richards, DL Hawksworth, and D Bass. 2011. Validation and justification of the phylum name Cryptomycota Phyl. Nov. *IMA Fungus* **2** (2): 173–75. https://doi.org/10.5598/imafungus.2011.02.02.08.

Kartali, T, I Nyilasi, B Szabó, S Kocsubé, R Patai, TF Polgár, G Nagy, C Vágvölgyi, and T Papp. 2019. Detection and molecular characterization of novel dsRNA viruses related to the Totiviridae family in *Umbelopsis ramanniana*. *Frontiers in Cellular and Infection Microbiology* **9** 1–13. https://doi.org/10.3389/fcimb.2019.00249.

Kitahara, R, Y Ikeda, H Shimura, C Masuta, and T Ezawa. 2014. A unique mitovirus from Glomeromycota, the phylum of arbuscular mycorrhizal fungi. *Archives of Virology* **159** (8): 2157–60. https://doi.org/10.1007/s00705-014-1999-1.

Koonin, EV. 2010. The wonder world of microbial viruses. *Expert Review of Anti-Infective Therapy* **8** (10): 1097–99. https://doi.org/10.1586/eri.10.96.

Lazarus, KL, and TY James. 2015. Surveying the biodiversity of the Cryptomycota using a targeted PCR approach." *Fungal Ecology* **14** 62–70. https://doi.org/10.1016/j.funeco.2014.11.004.

Lee, KM, WK Cho, J Yu, M Son, H Choi, K Min, YW Lee, and KH Kim. 2014. A comparison of transcriptional patterns and mycological phenotypes following infection of *Fusarium graminearum* by four mycoviruses. *PLoS ONE* **9**(6): e100989. https://doi.org/10.1371/journal.pone.0100989.

Li, P, S Wang, L Zhang, D Qiu, X Zhou, and L Guo. 2020. A tripartite ssDNA mycovirus from a plant pathogenic fungus is infectious as cloned DNA and purified virions. *Science Advances* **6** (14): 1–10. https://doi.org/10.1126/sciadv.aay9634.

Longcore, JE, AP Pessier, and DK Nichols. 1999. *Batrachochytrium dendrobatidis* Gen. et Sp. Nov., a chytrid pathogenic to amphibians. *Mycologia* **91** (2): 219–27. http://www.jstor.org/stable/3761366?origin=crossref.

Márquez, LM, and MJ Roossinck. 2012. Do persistent RNA viruses fit the trade-off hypothesis of virulence evolution? *Current Opinion in Virology* **2** (5): 556–60. https://doi.org/10.1016/j.coviro.2012.06.010.

Márquez, LM, RS Redman, RJ Rodriguez, and MJ Roossinck. 2007. A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science* **315** (5811): 513–15.

Marzano, SL, BD Nelson, O Ajayi-oyetunde, CA Bradley, TJ Hughes, and GL Hartman. 2016. Identification of diverse mycoviruses through metatranscriptomics. *Journal of Virology* **90** (15): 6846–63. https://doi.org/10.1128/JVI.00357-16.

Mondo, SJ, KH Toomer, JB Morton, Y Lekberg, and TE Pawlowska. 2012. Evolutionary stability in a 400-million-year-old heritable facultative mutualism. *Evolution* **66** (8): 2564–76. https://doi.org/10.1111/j.1558-5646.2012.01611.x.

Money, N. 2018. The rise of yeast: how sugar fungus shaped civilization. *Oxford University Press*. New York, NY.

Moore, D, GD Robson, and APJ Trinci. 2011. 21st century guidebook to the fungi. *Cambridge University Press*.

Neupane, A, C Feng, J Feng, A Kafle, H Bücking, and SL Marzano. 2018. Metatranscriptomic analysis and in silico approach identified mycoviruses in the arbuscular mycorrhizal fungus *Rhizophagus* Spp. *Viruses* **10** (12). https://doi.org/10.3390/v10120707.

Özkan, S, I Mohorianu, P Xu, T Dalmay, and RHA Coutts. 2017. Profile and functional analysis of small RNAs derived from *Aspergillus fumigatus* infected with double-stranded RNA mycoviruses. *BMC Genomics* **18** (1): 1–13. https://doi.org/10.1186/s12864-017-3773-8.

Papp, T, I Nyilasi, C Fekete, L Ferenczy, and C Vágvölgyi. 2001. Presence of double-stranded RNA and virus-like particles in *Rhizopus* isolates. *Canadian Journal of Microbiology* **47** (5): 443–47. https://doi.org/10.1139/cjm-47-5-443.

Pearson, MN, RE Beever, B Boine, and K Arthur. 2009. Mycoviruses of filamentous fungi and their relevance to plant pathology. *Molecular Plant Pathology* **10** (1): 115–28. https://doi.org/10.1111/j.1364-3703.2008.00503.x.

Powell, MJ. 2016. Handbook of the protists: Chytridiomycota. *Springer International Publishing*. JM Archibald *et al.* (*eds).*

Powell, MJ, PM Letcher, JE Longcore, and WH Blackwell. 2018. *Zopfochytrium* is a new genus in the Chytridiales with distinct zoospore ultrastructure. *Fungal Biology* **122** (11): 1041–49. https://doi.org/10.1016/j.funbio.2018.08.005.

Revill, PA, AD Davidson, and PJ Wright. 1998. Mushroom bacilliform virus RNA: the initiation of translation at the 5’ end of the genome and identification of the VPg. *Virology* **249** (2): 231–37.

Rigling, D, and S Prospero. 2018. *Cryphonectria parasitica*, the causal agent of chestnut blight: invasion history, population biology and disease control. *Molecular Plant Pathology* **19** (1): 7–20. https://doi.org/10.1111/mpp.12542.

Roossinck, MJ. 2011. The good viruses:viral mutualistic symbiosis. *Nature Reviews Microbiology.* **9,** 99-108.

Roossinck, MJ. 2015. Move over, bacteria! Viruses make their mark as mutualistic microbial symbionts. *Journal of Virology* **89** (13): 6532–35. https://doi.org/10.1128/jvi.02974-14.

Roossinck, MJ. 2005. Symbiosis versus competition in plant virus evolution. *Nature Reviews Microbiology* **3** (12): 917–24. https://doi.org/10.1038/nrmicro1285.

Roux, S, JR Brum, BE Dutilh, S Sunagawa, MB Duhaime, A Loy, BT Poulos, *et al.* 2016. Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **537** (7622): 689–93. https://doi.org/10.1038/nature19366.

Schmitt, MJ, and F Breinig. 2002. The viral killer system in yeast: from molecular biology to application. *FEMS Microbiology Reviews* **26** (3): 257–76. https://doi.org/10.1016/S0168-6445(02)00099-2.

Segers, GC, X Zhang, F Deng, Q Sun, and DL Nuss. 2007. Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. *Proceedings of the National Academy of Sciences of the United States of America* **104** (31): 12902–6.

Shahi, S, A Eusebio-Cope, H Kondo, BI Hillman, and N Suzuki. 2019. Investigation of host range of and host defense against a mitochondrially replicating mitovirus. *Journal of Virology* **93** (6): 1–15. https://doi.org/10.1128/jvi.01503-18.

Son, M, J Yu, and KH Kim. 2015. Five questions about mycoviruses. *PLoS Pathogens* **11** (11): 5–11. https://doi.org/10.1371/journal.ppat.1005172.

Spatafora, JW, Y Chang, GL Benny, K Lazarus, ME Smith, ML Berbee, G Bonito, *et al*. 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* **108** (5): 1028–46. https://doi.org/10.3852/16-042.

State of the world's fungi. 2018. Kew Royal Botanical Gardens. https://stateoftheworldsfungi.org/

Sun, Q, GH Choi, and DL Nuss. 2009. Hypovirus-responsive transcription factor gene Pro1 of the chestnut blight fungus *Cryphonectria parasitica* is required for female fertility, asexual spore development, and stable maintenance of hypovirus infection. *Eukaryotic Cell* **8** (3): 262–70. https://doi.org/10.1128/EC.00338-08.

Vaidya, JG & AS Rabba. 1993. Fungi in Folk Medicine. *Mycologist* **7**, 131-133

Vainio, EJ, J Jurvansuu, R Hyder, M Kashif, T Piri, T Tuomivirta, A Poimala, *et al*. 2017. *Heterobasidion* partitivirus 13 mediates severe growth debilitation and major alterations in the gene expression of a fungal forest pathogen. *Journal of Virology* **92** (5): 1–15. https://doi.org/10.1128/jvi.01744-17.

Wang, Shuangchao, Pengfei Li, Jingze Zhang, Dewen Qiu, and Lihua Guo. 2016. Generation of a high resolution map of SRNAs from *Fusarium graminearum* and analysis of responses to viral infection." Scientific Reports **6**: 1–12. https://doi.org/10.1038/srep26151.

Weinbauer, Markus G. 2004. Ecology of prokaryotic viruses. *FEMS Microbiology Reviews* **28** (2): 127–81. https://doi.org/10.1016/j.femsre.2003.08.001.

Weitz, J, and S Wilhelm. 2012. Ocean viruses and their effects on microbial communities and biogeochemical cycles. *F1000 Biology Reports* **8** (September): 2–9. https://doi.org/10.3410/B4-17.

Wilhelm, SW, and CA Suttle. 1999. Viruses and nutrient cycles in the sea aquatic food webs. *BioScience* **49** (October): 781–88. https://doi.org/10.2307/1313569.

Yu, X, B Li, Y Fu, D Jiang, SA Ghabrial, G Li, Y Peng, *et al*. 2010. A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proceedings of the National Academy of Sciences of the United States of America* **107** (18): 8387–92. https://doi.org/10.1073/pnas.0913535107.

Zhang, X, D Shi, and DL Nuss. 2012. Variations in hypovirus interactions with the fungal-host RNA-silencing antiviral-defense response. *Journal of Virology* **86** (23): 12933–39. https://doi.org/10.1128/jvi.00961-12.

Chapter 2 Survey of early-diverging lineages of fungi reveals abundant and diverse mycoviruses

Abstract

Mycoviruses are widespread and purportedly common throughout the fungal kingdom, although most are known from hosts in the two most recently diverged phyla, Ascomycota and Basidiomycota, together called Dikarya. To augment our knowledge of mycovirus prevalence and diversity in underexplored fungi, we conducted a large-scale survey of fungi in the earlier-diverging lineages, using both culture-based and transcriptome-mining approaches to search for RNA viruses. In total, 21.6% of 333 isolates were positive for RNA mycoviruses. This is a greater proportion than expected based on previous taxonomically broad mycovirus surveys and is suggestive of a strong phylogenetic component to mycoviral infection. Our newly found viral sequences are diverse, composed of dsRNA, positive-sense ssRNA, and negative-sense ssRNA genomes, and include novel lineages lacking representation in the public databases. These identified viruses could be classified into 2 orders, 5 families, and 5 genera, however half of the viruses remain taxonomically unassigned. Further, we identified a lineage of virus-like sequences in the genomes of members of Phycomycetaceae and Mortierellales that appear to be novel genes derived from integration of a viral RNA-dependent RNA polymerase gene. The two screening methods largely agreed in their detection of viruses; thus, we suggest that the culture-based assay is a cost-effective means to quickly assess whether a laboratory culture is virally infected. This study used culture collections and publicly available transcriptomes to demonstrate that mycoviruses are abundant in laboratory cultures of early-diverging fungal lineages. The

function and diversity of mycoviruses found herein will help guide future studies into mycovirus origins and ecological functions.

Importance

Viruses are key drivers of evolution and ecosystem function and are increasingly recognized as symbionts of fungi. Fungi in early-diverging lineages are widespread, ecologically important, and comprise the majority of the phylogenetic diversity of the kingdom. Viruses infecting early-diverging lineages of fungi have been almost entirely unstudied. In this study, we screened fungi for viruses by two alternative approaches: a classic culture-based method and by transcriptome-mining. The results of our large-scale survey demonstrate that early-diverging lineages have higher infection rates than have been previously reported in other fungal taxa, and that laboratory strains worldwide are host to infections, the implications of which are unknown. The function and diversity of mycoviruses found in these basal fungal lineages will help guide future studies into mycovirus origins and their evolutionary ramifications and ecological impacts.

Introduction

Fungal viruses (mycoviruses) have been reported from all major fungal taxonomic groups (Ghabrial 2015; Son 2015). However, the overwhelming majority of mycoviruses have been identified in hosts belonging to just two phyla—Ascomycota and Basidiomycota (known together as "Dikarya")—though Kingdom Fungi is comprised of at least 8 phyla (James 2020; Spatafora 2017; Hibbett 2007) (Figure 2.1). Known mycovirus infections in non-Dikarya are limited to an early report of virus-like particles from ultrastructure studies of *Allomyces arbusculus* (Khandjian 1974), and sequence-based identification in: *Rhizopus oryzae* (Papp 2001), six isolates of arbuscular mycorrhizal fungi (Kitahara 2014; Neupane 2018), *Umbelopsis ramanniana* (Kartali 2019), *Entomophthora muscae* (Coyle 2018; Nibert 2019), and *Rhizopus*

microsporus (Espino-Vázquez 2020). While it has become clear that mycoviruses are widespread, disproportionate sampling across the fungal kingdom has resulted in an incomplete understanding of mycovirus diversity and prevalence. By sampling unexplored and phylogenetically diverse fungal lineages, we predicted to find equally diverse viruses that could empower meaningful inquiries into the origins and ecological functions of mycoviruses.

The conventional approach to mycovirus discovery exploits the observation that double-stranded RNA (dsRNA) genomes predominate among known mycoviruses (Ghabrial 2015). The most common method uses cellulose chromatography to purify dsRNA from total RNA extracts, and effectively isolate mycoviruses with dsRNA genomes and the dsRNA replicative intermediates of single-stranded RNA genomes (Morris & Dodds 1979). This approach is quick and inexpensive; however, its exclusive use ignores DNA viruses and reinforces the bias that mycoviral genomes are predominantly composed of RNA. Furthermore, viral RNA is detected visually by gel electrophoresis, which could result in false negatives in instances of low-titer infections. To our knowledge, estimates of the proportion of false negatives based on chromatography have not been reported. We tested the relative accuracy of cellulose chromatography for identifying fungal isolates with viral infection and compared to more indirect sequence-based alternative approaches.

In silico approaches to mycovirus discovery have become more common (Nerva 2015; Marzano 2016; Gilbert 2019). These approaches search transcriptomic data for sequences similar to RNA-dependent RNA polymerase (RdRp), which is the only conserved gene among RNA viruses and is diagnostic of infection. Conceptually, these approaches should be more sensitive to low titer viral infections because they do not rely on gel visualization and high transcriptome sequencing depth is readily achieved. Indeed, RNA sequencing has revealed cryptic co-infection

in isolates previously thought to be singularly infected (Osaki 2016; Myers, personal observation). As sequencing costs have dropped, the availability of RNA-Seq data has grown, which presents a novel opportunity to probe existing datasets to characterize viruses for the first time. Gilbert *et al*. (2019) developed a data-mining pipeline for mycovirus identification, applied it to public transcriptomes of fungi in the sub-phylum Pezizomycotina, and discovered 52 novel mycoviruses, demonstrating the utility of this approach. However, such approaches have yet to be applied widely to non-Dikarya lineages.

Our primary goal was to improve understanding of prevalence and sequence diversity of RNA mycoviruses in the early-diverging lineages of Fungi, specifically Mucoromycota, Zoopagomycota, Chytridiomycota, Blastocladiomycota, Neocallimastigomycota, and Cryptomycota/Rozellomycota. We screened early-diverging lineages with both a culture-based chromatography approach ("*in vitro*") and a transcriptome data-mining approach ("*in silico*"). When possible, we compared methods by screening the same isolates used for transcriptome-generation with the *in vitro* method. In total, we screened 333 hosts, of which 72 (21.6%) harbored viruses. These results demonstrate that mycoviruses are abundant in the early-diverging lineages of the fungal kingdom, including in research laboratory cultures around the world.

Methods

In vitro screening

Fungal cultures

Cultures were obtained from the Collection of Zoosporic Eufungi at the University of Michigan (CZEUM) (recently founded from the Joyce Longcore U. Maine Collection (JEL) and the University of Alabama Chytrid Culture Collection (UACCC)), Agricultural Research Service Culture Collection (NRRL), Collection of Entomopathogenic Fungal Cultures (ARSEF), and the

collections of the authors (Supplementary Table 2.1). We grew isolates in media appropriate for their nutritional needs until sufficient biomass accumulated (3 days–2 weeks depending on the species), harvested tissue, and ground it by sterile mortar and pestle in liquid nitrogen. With every batch of fungi screened, we harvested and processed a mycovirus-infected strain of Ustilago maydis as a positive control for degradation of mycoviruses by RNases.

Column preparation and dsRNA extraction

We screened cultures for RNA mycoviruses by dsRNA extraction and purification by cellulose chromatography as described by Okada *et al*. 2015 with slight modifications. Before RNA extraction, we prepared columns by piercing the bottom of a 0.5 mL tube with an 18-gauge needle, packing it with approximately 90 mg cellulose D powder (Advantec), placing it in a 2 mL microcentrifuge tube, and adding 400 µL of freshly prepared 1X Sodium Chloride-Tris-EDTA (STE) with 16% ethanol. Immediately before use, we centrifuged the columns briefly and discarded the flow-through. We extracted RNA by adding one mL of TRIzol reagent (Invitrogen) to approximately 0.5 mg finely ground frozen tissue and either freezing at -20°C for later processing or incubating for 10 minutes at room temperature, adding 200 µL of chloroform, mixing by inversion, incubating for three minutes at room temperature, and centrifuging at 12 000 x g at 4°C for 15 minutes. We made a 16% ethanol solution with the supernatant, added it to the cellulose column, collected and discarded the flow-through, washed the column three times with 400 µL of 1X STE with 16% ethanol, thoroughly dried it by centrifugation, and eluted columns with 400 µL 1X STE. We precipitated dsRNA by adding 40 µL 3 M sodium acetate and 1 mL ethanol, centrifuged at 15 000 x g for 5 minutes, pipetted off the supernatant, and allowed the tubes to dry before reconstitution with water. We treated samples with S1 Nuclease and

DNase 1 per manufacturer instruction before visualizing by agarose gel electrophoresis. We considered samples positive if a band was visible (Figure 2.2).

In silico screening

We obtained unassembled RNA-Seq data from the SRA database (Supplementary Table 2.2), stringently filtered raw reads for quality (min phred score 20) using the fastxtoolkit (Hannon 2010), and assembled contigs de novo with Trinity assembler (Grabherr 2011). We predicted ORFs with Transdecoder (Haas 2013) using default parameters and queried the protein translations with hmmscan (Wheeler 2013) against a custom RdRp HMM database. We constructed the RdRp database as in Gilbert 2019: we downloaded entire alignments of Pfam families RdRp_1, RdRp_2, RdRp_3, RdRp_4, and Mito_RdRp, and generated HMM profiles from each using hmmbuild (HMMER2; hmmer.org). We further queried each RdRp profile hit found in the transcriptome ORFs with TBLASTN and BLASTP to the NCBI nt and nr databases (downloaded 09/18/2019), respectively, and considered isolates positive for viral infection if the resultant hits were viral sequences with e-value < e-10. To ensure the virus was exogenous to the host genome, we blasted viral contigs against the hosts' genome when available, or, if unavailable, the genome of the closest relative.

To improve assemblies, we used the nucleotide sequence of the Trinity-assembled viral contig as the seed for contig extension with PRICE (Ruby 2013) with parameters of minimum 30 nt overlap for mini-assembly, minimum 80% identity for contig-edge assembly, 90% identity to starting contigs, and for 10 cycles, and using loosely-filtered raw reads (min phred score 5) (MacManes 2014). This most often resulted in contigs that were unchanged after 10 cycles, which we considered complete. If contigs were updated in the procedure, we ran 10 additional cycles with the updated contig as the starting seed.

As a final quality control check of our virus genomes, we re-assembled the viromes of a subset of isolates by first aligning loosely-filtered raw reads (min phred score 5) to a reference genome using STAR (Dobin 2013), assembled unmapped reads de novo using Trinity, and continued the pipeline exactly as described above. In all cases at least one viral contig was extended by this method but the RdRp region was most often unchanged. For additional informatics details and code see https://github.com/jimyers/Mycoviruses-in-early-diverging-fungal-lineages.

Mycovirus sequencing

Pacific Biosciences sequencing

We prepared purified dsRNAs obtained from *in vitro* screens of *Allomyces* sp. JMM01, *Allomyces* sp. DJ-02, *Allomyces* sp. DJ-07, *Zopfochytrium polystomum* WB228, and *Ustilago maydis* (as a control) as described in Roossinck 2010 with slight modifications. Briefly, we purified dsRNAs by cellulose chromatography as described above, except the final elution was performed with 20 µL 2X STE, pH 8.0. We mixed 1 µL dsRNA with 7 µL H2O and 2 µL tagged random 12-mer at 20 µM (5'ACCTTCGGATCCTCC N12 3'), placed the tube in boiling water for two minutes to melt strands, immediately quenched on ice, then used the Omniscript reverse transcription kit per manufacturer protocol, incubated tubes on ice for 10 minutes followed by 60 minutes at 50° C , removed unreacted template with 2 µL Ribonuclease A (5 mg/mL) and incubated at room temperature for 15 minutes followed 3 minutes at 80° C to denature enzymes. We used a Qiagen PCR purification kit per manufacturer protocols to clean up PCR products and amplified them by PCR with individually barcoded primers using GoTaq polymerase followed by gel-extraction using a Qiagen gel extraction kit and manufacturer's protocol. For each isolate, we sequenced 120 ng of product on the PacBio RSII at the University of Michigan Advanced

Genomics Core (UM-AGC). We analyzed reads with SMRT Portal (parameters: minimum barcode score 22, minimum 5 passes, minimum 90% accuracy) and removed adapters with cutadapt followed by read correction, trimming, and assembly with Canu (v1.3) (parameters: minOverlapLength=100, minReadLength=150, errorRate=0.035, estimated genome size of 10,000 bp). For quality control, we compared a *U. maydis* contig to the GenBank *U. maydis* H1 cap-pol fusion nucleotide sequence (NC_003823; fungal isolate unknown) which matched with 99% query cover and 93% identity.

(ds)RNA-Seq and total RNA-Seq

We submitted purified dsRNAs from *in vitro* screens of *Cladochytrium* sp. JEL861, *Allomyces arbusculus* North Carolina 2, *Umbelopsis nana* TLT 204, and *U. maydis* (as a control) to the UM-AGC for 2x150 sequencing on Illumina MiSeq with the following modifications to the manufacturer's instructions for library preparation as in Sasai *et al*. 2018: fragmentation was conducted with 87.5 ng of purified dsRNA for 20 minutes at 94º C in First Strand Synthesis buffer with random primers and the first strand synthesis reaction was conducted for 30 minutes at 42º C. We submitted total RNA extractions of *Allomyces* sp. JMM01 and *Z. polystomum* WB228 to the UM-AGC for library preparation using the Illumina TruSeq stranded mRNA protocol and 2x50 sequencing using an Illumina HiSeq 4000. All RNA-seq was processed to remove adapters and low-quality sequences from paired-end data using trimgalore (min quality threshold 5) (https://github.com/FelixKrueger/TrimGalore), followed by assembly de novo with Trinity. We performed ORF prediction, HMM queries for RdRp homologs, and PRICE extension as described above.

For transcriptomes sequenced by JGI and not published previously (Supplementary Table 2.2), stranded cDNA libraries were generated using the Illumina Truseq Stranded RNA LT kit.

mRNA was purified from 1 µg of total RNA using magnetic beads containing poly-T oligos, fragmented and reverse-transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 8-10 cycles of PCR. The prepared library was quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument, multiplexed with other libraries, and the pool of libraries was then sequenced on the Illumina platform (HiSeq 2000/2500 or NovaSeq) following a 2x150 indexed run recipe.

Comparison of screening approaches

One isolate, *Mortierella humilis* PMI 1414, produced negative results by *in vitro* virus screen but positive results by the *in silico* method. We designed virus-specific primers based on the *in silico* results and conducted reverse-transcription PCR using SuperScript IV reverse transcriptase (ThermoFisher) per manufacturer protocol. For a proxy of viral titer, we calculated the abundance of viral transcripts generated in the transcriptome with the align_and_estimate_abundance.pl script of the Trinity package (https://github.com/trinityrnaseq/trinityrnaseq), employing bowtie for alignment and the RSEM estimation method.

Mycovirus sequence analysis and phylogenetics

We classified each new mycoviral sequence by top blast hit and grouped them into clusters corresponding to "branches" described in the most recent RNA virus phylogeny (Wolf 2018). Open-reading frames including each RdRp gene were predicted and translated using the universal genetic code except for sequences with top blast hits to mitoviruses, for which we used the standard mitochondrial genetic code.

Despite our best attempts to resolve assemblies, some remained fragmented. To avoid over-reporting, we applied conservative criteria to limit the number of viruses reported for each isolate in a biologically relevant manner. The criteria for inclusion in our analyses were as follows: if only one viral contig was identified per strain per branch, we included it. For branches with more than one viral contig per isolate, all contigs with > 60% coverage of the RdRp conserved domain were included, or, if all contigs for that branch contain <60% of the RdRp, then we kept the contig with highest coverage. Three isolates had only two viral contigs assembled, one of which contained the C-terminus of the RdRp and the other contained the N-terminus (*B. trispora*, *M. verticillata*, *R. intraradices*). In these three instances, we concatenated the two contigs.

For each branch, we inferred RdRp gene trees with viral contigs that met our criteria, their top BLAST hits, and reference sequences (Supplementary Table 2.3). A sixth tree (Figure 2.9) includes sequences with BLAST similarities to viruses currently unassigned by viral taxonomy. We aligned sequences with MAFFT version 7 using the E-INS-i algorithm (Katoh 2013) and trimmed the resulting alignments using the -automated1 method in TrimAl (Capella-Gutiérrez 2009). We determined the best-fit model of amino acid substitution for each alignment with ProtTest 3.4 (Darriba 2011) and reconstructed trees with the maximum likelihood approach implemented in RaxML by the rapid bootstrap analysis (-f a) with 100 replicates (Stamatakis 2014).

Results and Discussion

Viral prevalence

Through deep sequencing approaches, recent studies have revealed the abundance of viruses in pathogenic fungal species, marine environments, and soils (Marzano 2016; Nerva 2015; Sutela 2019). We build on this momentum by exploring viruses hosted in species from the deep branches of Kingdom Fungi. In doing so, we more completely characterized the abundance of viruses across the Kingdom and uncovered novel viral sequence diversity linked to specific fungal hosts. We determined the first sequences of mycoviruses in the particularly under-explored phyla Blastocladiomycota, Neocallimastigomycota, and Chytridiomycota (Figure 2.1). In total, we screened 333 fungi spanning six phyla by either or both *in vitro* (cellulose chromatography) and *in silico* (transcriptome-mining) methods (Supplementary Table 2.1) and found one or more viruses in 72 isolates (21.6%), 65 of them not previously known as mycoviral hosts (Figure 2.1). Of the 36 hosts for which we obtained viral sequence data —either by assembling from host transcriptomes or direct sequencing—we generated 154 complete or partial viral sequences which, using the criteria described above, we conservatively reduced to 85 unique viruses.

All sampled phyla were hosts to viruses except the poorly-sampled Cryptomycota. Infection prevalence at the phylum level ranged from 15.9–40.0%, and 0¬89% at the sub-phylum and order levels, but was highest in Glomeromycotina, Entomophthoromycotina, and Cladochytriales (Table 2.1, Figure 2.3). Glomeromycotina are also notorious hosts of endosymbiotic bacteria which may have originated by the invasion of free-living bacteria following hyphal damage by herbivores or other fungi (Bonfante and Desirò 2017); it is possible that mycoviruses similarly originated in these fungi by invasion through damaged cell walls. In

the zoosporic fungi, we predicted that polycentric organisms would be disproportionately virally infected since this mode should support viral transmission more frequently than monocentric growth. In monocentric development a zoospore encysts on a substrate and develops into a single zoosporangium—the structure which will ultimately cleave into many zoospores. Polycentric development, on the other hand, involves a single zoospore producing multiple zoosporangia with cytoplasmic continuity via aseptate rhizomycelium. Thus, a single viral infection in a polycentric organism could lead to many spores being infected while monocentric development may allow for clearing of viral infection through selection between zoospores. Within the Chytridiomycota, the percent of polycentric organisms that were viral positive (46.2%, n=26) was significantly greater ($P<0.00004$, Fisher's exact test) than monocentric organisms (9.2%, n=120). Ten of the twelve positive polycentric isolates were from order Cladochytriales, however, and consequently we were unable to determine whether this trend is a result of the morphological trait or susceptibility traits that track the host phylogeny without additional data from non-Cladochytrialean polycentric organisms. Although the sample size is inadequate for definitive results regarding Neocallimastigomycota, viral infection did not favor polycentric over monocentric organisms in this phylum (polycentric n=4, monocentric n=4, 25% of each was viral positive).

Early in the study of mycoviruses, Bozarth (1972) estimated 10–15% infection prevalence for fungal cultures predominantly made up of Dikaryotic lineages but including some organisms in the Mucoromycota and Blastocladiomycota. A recent *in silico* survey of mycoviruses in Pezizomycotina (Ascomycota) revealed infection prevalence ranging up to 50% in some classes, but an overall rate of about 8% (47/569) for the subphylum (Gilbert 2019). Our estimate is approximately 22% (Table 2.1) across the early-diverging phyla. Interestingly, we

also found substantial variation across taxa (Table 2.1, Figure 2.3). Among the phyla with the highest infection rates, Mucoromycota and Zoopagomycota, viral prevalence at the subphylum level ranged from 19.7-88.9% (Table 2.1), a surprising difference from the 8% prevalence found for the same taxonomic level, Pezizomycotina, in Ascomycota. These datasets were not explicitly controlled for geographic origin of isolates, which could influence prevalence if, for example, some geographic regions had a higher proportion of viruses relative to others. However, the isolates in our study predominantly originate from distinct populations across the globe (Supplementary Table 2.1). Our results indicate higher prevalence in some subphyla of basal fungal lineages compared to Pezizomycotina. From our comparison of viral prevalence at the subphylum level between Pezizomycotina (Ascomycota) versus the early-divering lineages, it is tempting to speculate on differing viral prevalence at the phylum level. However, a taxonomically thorough and geographically controlled study of viral prevalence in the other subphyla of Ascomycota and in Basidiomycota is needed in order to make a direct comparison. Additionally, anecdotal evidence suggests mycoviruses are commonly lost through culturing. Our estimates from the *in vitro* assays are, thus, perhaps underestimates of the true infection load in nature. Nonetheless, it is compelling to consider the implications of years or decades of stable maintenance of mycoviruses in culture on how the viruses affect the fungi. Future studies are needed to better characterize the effects of these mycoviruses on these phylogenetically diverse hosts.

Taken together, these findings indicate a strong phylogenetic component to viral prevalence. The earlier-diverging fungal lineages surveyed in this study typically share the trait of coenocytic mycelia, which may benefit the transmission of obligate symbionts and thus

facilitate mycoviral infection; if true, it is possible that septa evolved, at least in part, as a viral defense mechanism to limit the spread of viruses throughout a mycelium.

Comparison of screening approaches

The results of screening the same isolate by chromatography and transcriptome-mining were mostly in agreement (*i.e.* at least one band of dsRNA was present by chromatography and RdRp sequence(s) were identified *in silico* or no dsRNA was present by chromatography and no RdRp sequences were identified) (85.7%, n=21). Of the three isolates where the two methods varied in detection, two were found positive *in vitro* but negative *in silico*, and one with the inverse result. For *Mortierella humilis* (PMI 1414), where the *in silico* method revealed viral sequences but the *in vitro* method did not, RT-PCR confirmed viral presence in our culture. From the transcriptome data, we determined the virally-derived sequences made up 2.58 transcripts-per-million (TPM). If we consider this a proxy of viral titer, the low viral abundance in the host likely accounts for the initial negative result obtained with the *in vitro* method. From *Umbelopsis nana* TLT204, which was positive by in-house chromatography but negative *in silico*, we sequenced purified dsRNA and assembled a toti-like viral contig containing a complete RdRp domain and an L-A virus major coat-protein domain, confirming viral presence in our culture. Most likely, the virus was lost through subculture of this isolate in the laboratory in which transcriptome analysis was performed.

Paired comparison of screening approaches suggests minimal discrepancies, but overall more viruses were detected by the *in silico* method (33.3% positive) than *in vitro* (17% positive). The *in silico* method is likely to be more sensitive than the *in vitro* method, but it is possible that the higher virus detection rate *in silico* was a result of phylogenetic skews of the data sets rather than methodology. Approximately 73% of the cultures screened *in vitro* were in the

Chytridiomycota or Blastocladiomycota—the lesser-infected phyla. Accordingly, 64% of the isolates screened *in silico* were in the Mucoromycota or Zoopagomycota—the greater-infected phyla. Thus, we recommend screening by transcriptome-mining when the data is available, but the chromatography approach is a good alternative for very cost-effective initial screening.

Mycoviral diversity

We found mycoviruses in all five "branches" of the Riboviria (RNA virus) tree published by Wolf et al. (2018) (Figures 2.4–2.8). Many of these newly found viruses will need to be described as new taxa at the levels of genus, family, and order. Additionally, some viruses, including a group of related, endogenous viral-like sequences, represent novel diversity such that they were unable to be assigned to a branch, though they show sequence similarity to other unplaced mycoviruses (Figure 2.9). Coinfection of a single host by multiple viruses was very common (mean and median number of viruses per host = 2.4, 1.5, respectively). Even after filtering by our conservative criteria, a single host, *Kickxella alabastrina* (Kickxellomycotina), was found to harbor at least 11 unique viruses. Such a result would be unsurprising in macrobes, such as humans, but is staggering for a microfungus.

Riboviria: "Branch 1"

Branch 1 (Figure 2.4) is composed of positive-sense ssRNA viruses including the bacteria-infecting *Leviviridae*, the mitochondria-infecting mitoviruses (*Narnaviridae*), and the cytoplasmic narnaviruses (*Narnaviridae*) and ourmiaviruses (*Botourmiaviridae*). In total, 22 viruses were assigned to branch 1 based on blast similarity. A complete RdRp gene from the transcriptome of *Chaetocladium brefeldii* was most similar to viruses in family *Leviviridae*. These viruses are currently only known from bacteria, so it is possible, perhaps likely, that this contig derives from the viruses of a bacterial endosymbiont of *C. brefeldii*.

Interestingly, the new viruses in this group are mostly hosted by fungi in the Mucoromycota or Zoopagomycota. Previous evidence suggests the origin of mitoviruses was a common ancestor of Mucoromyota and Zoopagomycota, which was followed by cospeciation of host and mitoviruses (Nibert 2019) and horizontal transmission to plants (Nibert 2018). Our findings support this hypothesis, but sequence data is still limited for viruses in Chytridiomycota, Blastocladiomycota, and Neocallimastigomycota. Additional sequencing of the viral positives in those groups found by chromatography in this study will more strongly support or refute this hypothesis.

Riboviria: "Branch 2"

The picornavirus supergroup makes up branch 2 (Figure 2.5), which also includes *Partitiviridae*, *Amalgaviridae*, *Barnaviridae*, and *Potyviridae*. In total, 9 viruses were assigned to branch 2 based on BLAST similarity. The hosts of these contigs represent the Mucoromycota, Zoopagomycota, Neocallimastigomycota, and Chytridiomycota. A virus of *Entomophthora muscae* nested within the picorna-like viruses, specifically in the genus *Iflavirus*, which is only known to infect insects. Thus, it may have originated by cross-kingdom horizontal transfer. A virus of the arbuscular mycorrhizal fungus *Rhizophagus irregularis* (Glomeromycotina) is nested within the plant-infecting potyviruses and likely also arose by horizontal transmission given the tight mutualism and nutrient exchange between *R. irregularis* and its plant hosts.

Riboviria: "Branch 3"

Branch 3 (Figure 2.6) includes the alphavirus and flavivirus supergroups. In total, 6 contigs were assigned to this branch based on BLAST similarity. We found 3 new tombus-like viruses, hosted by *Mortierella selenospora* (Mucoromycota), *Syncephalis fuscata* (Zoopagomycota), and *Anaeromyces* sp. (Neocallimastigomycota). Two arbuscular mycorrhizal

fungi, *R. irregularis* and *Paraglomus brasilianum*, contained viruses related to the predominantly plant-hosted viruses in *Virgaviridae*. Another Glomeromycotina fungus, *Geosiphon pyriformis*, was host to a tymo-like virus positioned basally in the order *Tymovirales* and may represent a new family in this order.

Riboviria: "Branch 4"

The largest number of new contigs were assigned to branch 4 (Figure 2.7), which includes *Totiviridae*, *Chrysoviridae*, *Quadriviridae*, and *Reoviridae*. In total, our analysis assigned 41 contigs to this group. Hosts are largely represented by Chytridiomycota and Blastocladiomycota, including substantial co-infection within the same individual, but also by Mucoromycota and Zoopagomycota.

Ten of these viruses group within the genus *Totivirus*, and three new viruses grouped within the genus *Victorivirus*, both genera belonging to the family *Totiviridae*. By our analysis, these two genera are composed of fungal viruses truly spanning the Kingdom, including hosts in Chytridiomycota, Blastocladiomycota, Zoopagomycota, Mucoromycota, Ascomycota, and Basidiomycota. This finding may align with the cospeciation hypothesis for this viral family: Göker *et al*. (2011) presented evidence that *Totiviridae* speciate through codivergence with their hosts. The predominance of totiviridae-like sequences in early-diverging lineages supports this hypothesis and suggests that an early fungal ancestor harbored ancestors of *Totiviridae*. However, this origin hypothesis does not preclude occasional horizontal transfer, many instances of which are suggested by our phylogenetic tree. Further, this viral family also include two genera, *Leishmaniavirus* and *Giardiavirus*, that infect protozoa; since protozoans and fungi are polyphyletic, ancient horizontal transfer between host groups is likely.

Basal to *Victorivirus*, a clade formed including viruses from *K. alabastrina*, *Operculomyces laminatus*, the ampicomplexan *Eimeria tenella*, and multiple diatom-colonies. This finding aligned with our prediction of close-relatedness between formerly reported diatom-associated viruses and the newly reported sequences from Chytridiomycota, a group that contains diatom parasites, but lacked bootstrap support. Viruses from multiple isolates of *Allomyces* sp. form a clade sister to *Botybirnavirus*, likely a new genus. Generally, many of the new sequences in this branch represent novel diversity that will require deeper viral characterization to phylogenetically classify with confidence.

Riboviria: "Branch 5"

We identified just one viral contig in branch 5, which includes the negative-sense ssRNA viruses *Mononegavirales* and *Bunyavirales* (Figure 2.8). The virus found in *Mortierella minutissima* is most closely related to the under-sampled yueviruses known from insect hosts. Viruses with fungal hosts have not previously been reported from this, or related, lineages and the *M. minutissima* virus may represent an undescribed family.

Riboviria: Unassigned

Some viral contigs either did not have significant BLAST hits or hit to known viruses currently unassigned to a taxonomic grouping (n=12) (Figure 2.9). These new viral contigs appear to be related to unplaced "bipartite mycoviruses" and *Curvularia thermotolance virus*, known to form a tri-partite mutualism with its fungal host and a plant (Marquez 2007). Among these viruses we uncovered is a novel lineage found endogenous in hosts' genomes. Intriguingly, we identified a highly supported clade of virus-like sequences only in the genomes of fungi in the Mucoromycota including two species of *Phycomyces*, *Dissophora ornata*, *Lobosporangium transversale*, and multiple *Mortierella* spp. (Figure 2.9). Conserved RdRp domains were

identifiable in all 9 of these sequences; all lacked specific hits to reverse transcriptome domains and thus are not believed to be retroviruses. Rather, the phylogenetic conservation of these endogenous virus-like sequences suggests that, before the divergence of Mucorales and Mortierellales, an endogenization event, in which a viral gene was reverse-transcribed and integrated into the host genome, occurred that was conserved in members of Phycomycetaceae and Mortierellales. Integration of viral genes into host genomes is a well-known phenomenon generally, and the transfer of genes from dsRNA viruses to Fungi has been reported by others (Frank and Wolf 2009; Taylor and Bruenn 2009; Liu 2010). We identified the new sequences via their mRNA transcripts with our *in silico* approach, which confirms their expression. Thus, this previously unreported instance of viral endogenization seemingly resulted in a novel fungal gene, the function of which remains to be tested.

Conclusions

The results of this study provide a new perspective on mycoviral prevalence in fungi and the phylogenetic diversity of viruses. The results also aid in the identification of hosts of environmentally derived virus samples. For example, many of our mitovirus-like viral contigs appear to be related to *Mitovirus* species from soils; we suggest that these soil-derived mitoviruses are likely hosted by fungi in Mucoromycota.

The ecological implications of viruses in these deep branches of the fungal tree are currently only a matter of speculation, but their role in natural ecosystems may be of great importance. The host organisms studied embody a broad diversity of ecological niches including saprotrophs, plant mutualists, obligate and opportunistic pathogens, and parasites of plants, invertebrates, animals, protists, and other fungi. Even slight effects on the growth rate of saprotrophs, for instance, could have significant impacts to nutrient cycling on the global scale.

By searching unexplored and under-explored fungal lineages we uncovered novel mycoviral diversity and discovered that fungal viruses are indeed ubiquitous throughout the fungal kingdom, detected now in nearly every phylum. Our data suggest that early-diverging lineages may harbor greater viral prevalence than the Dikarya, but there is wide variation across lineages. A caveat of comparisons across major taxonomic groups is that without broad and deliberate taxonomic and geographic sampling, comparison of rates of infection are subject to high error in estimation. Further, by searching publicly available transcriptomes as well as cultures from collections that are distributed to researchers globally, we learned that mycoviruses are abundant in research organisms used in laboratories worldwide. Mycoviruses are known to be persistent and often asymptomatic, but also to cause variable phenotypic alterations such as in pigmentation, growth rate, and virulence. The implications for cryptic mycoviral infection in laboratory cultures are currently unknown but provide a guide for future studies into mycovirus origins and ecological functions.

Acknowledgments

The authors thank Georgianna May and Kerry O'Donnell for providing fungal cultures, and Jesús Peña at UC Riverside and Sabrina Blakenship and Jeff Beard at Oregon State University for efforts to generate transcriptomes and genomes for the Zygolife project. Thanks to Andreas Brachman for depositing RNA-Seq data of *Geosiphon pyriforme* into GenBank. We thank Kerrie Barry and Chris Daum and his team for coordination and sequencing of the transcriptomes generated by JGI and used in this work. We thank Melissa Duhaime and Alisha Quandt for helpful conversation.

Author Contributions

The study was conceived of by JMM and TYJ. *In vitro* screenings were performed by JMM, RAC, and NT. Cultures were provided by JEL and GB. Cultures were grown by JMM, AEB, DRS, NT, and RAC. *In silico* screenings, sequencing, analyses, and manuscript preparation was performed by JMM. Transcriptomes were prepared and/or contributed by JMM, DRS, GB, JES, DCH, JO, PL, JWS, YC, LC, AMD, AG, IVG, and TYJ. All authors provided helpful comments on the manuscript.

Funding

This work was supported by US National Science Foundation (NSF) grants: DEB 1929738 to TYJ, DEB 1737898 to GB, DEB 1441715 and DEB 1557110 to JES, DEB 1441604 to JWS, and NSF fellowship DGE 1256260 to JMM. The work conducted by the U.S. Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

Data Availability Statement

The sequences generated for this study can be found in genbank BioProject ######. Sequence alignments, HMMs, and code can be found on github at https://github.com/jimyers/Mycoviruses-in-early-diverging-fungal-lineages

Figures

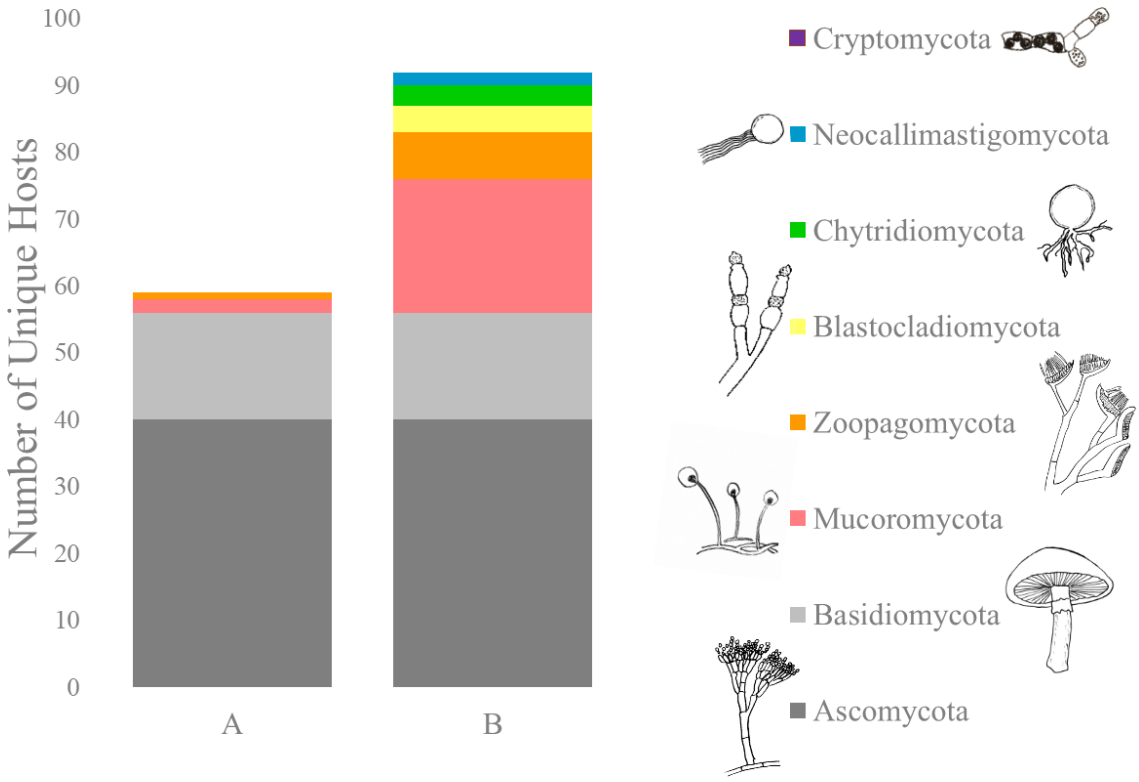


Figure 2.1 Barplot showing the number of unique hosts of exogenous mycoviruses, per phylum, as represented in GenBank before this study (A) and with the data from this study added (B).

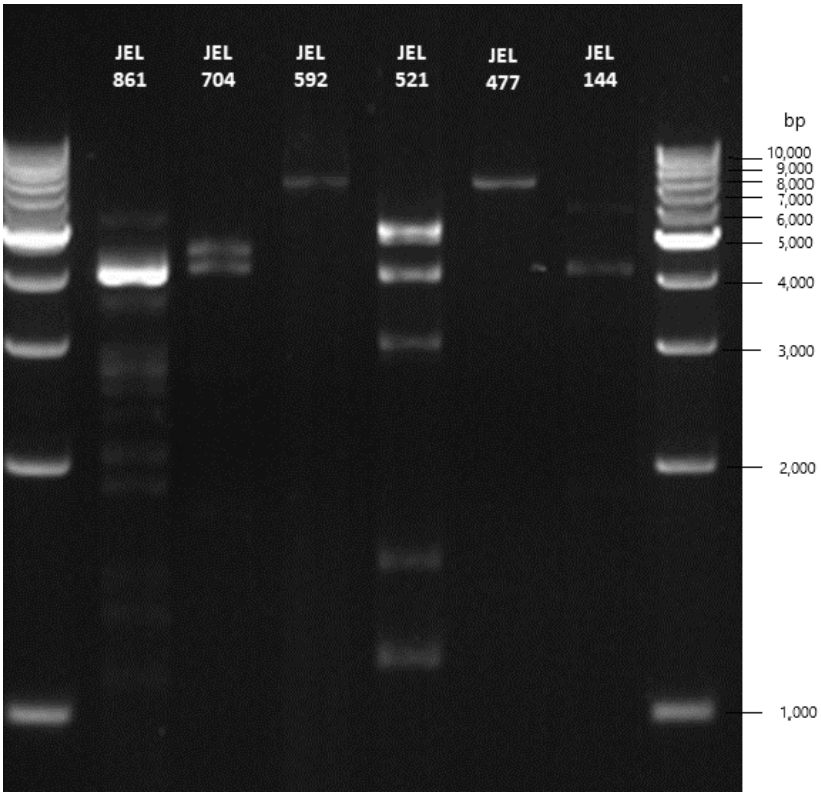


Figure 2.2 Agarose-gel electrophoresis image of the purified dsRNAs of six isolates of *Cladochytrium* sp. with varied banding patterns, flanked by 1kb ladders. The first sample (lane 2) shows dsRNA of *Cladochytrium* sp. JEL861, which was sequenced in this study.

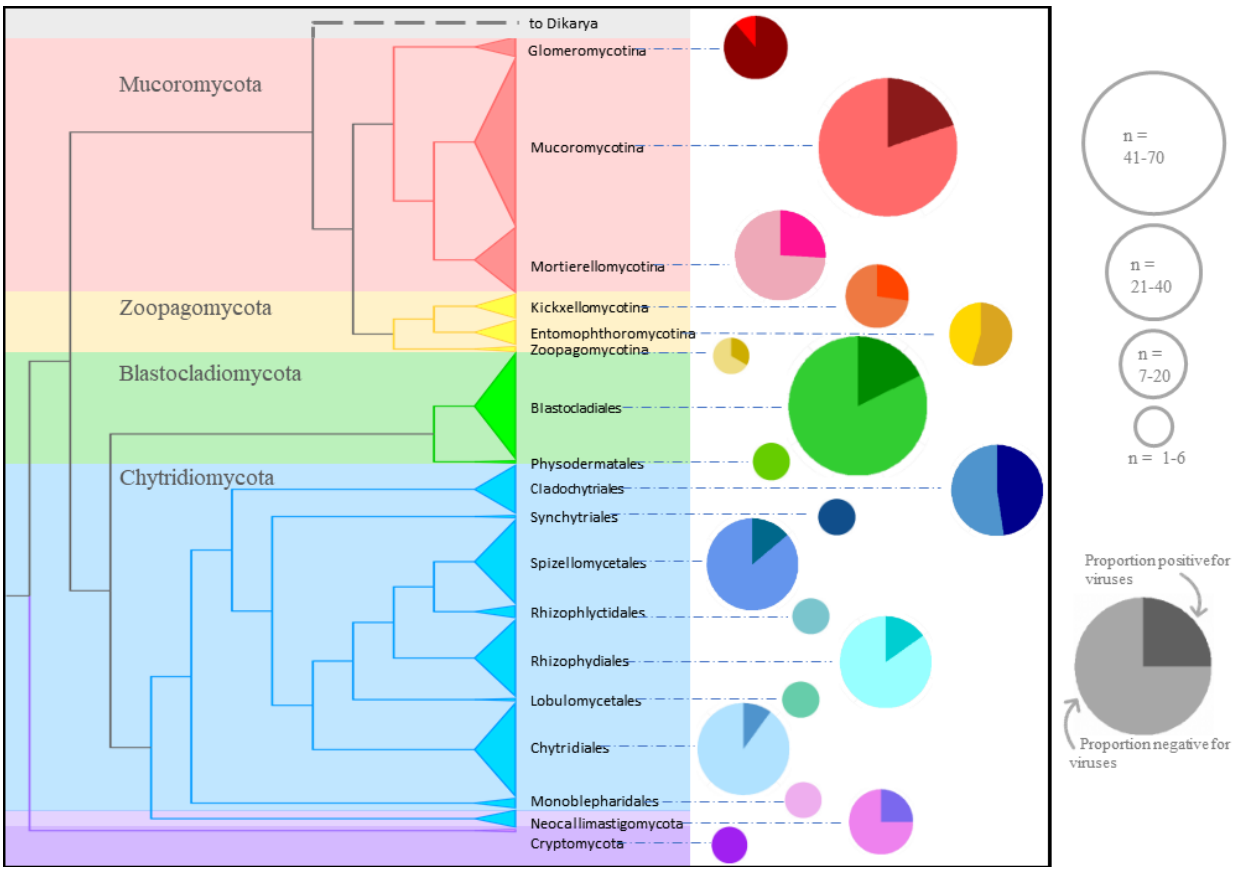


Figure 2.3 Cladogram of the organisms screened for viruses by both methods in this study. Size of the collapsed clades is proportional to number of isolates screened. Pie charts indicate the proportion of isolates in each taxon that were viral positive (darker shade); pie charts are sized according to number of isolates screened. Whole pies had a 0% infection rate.

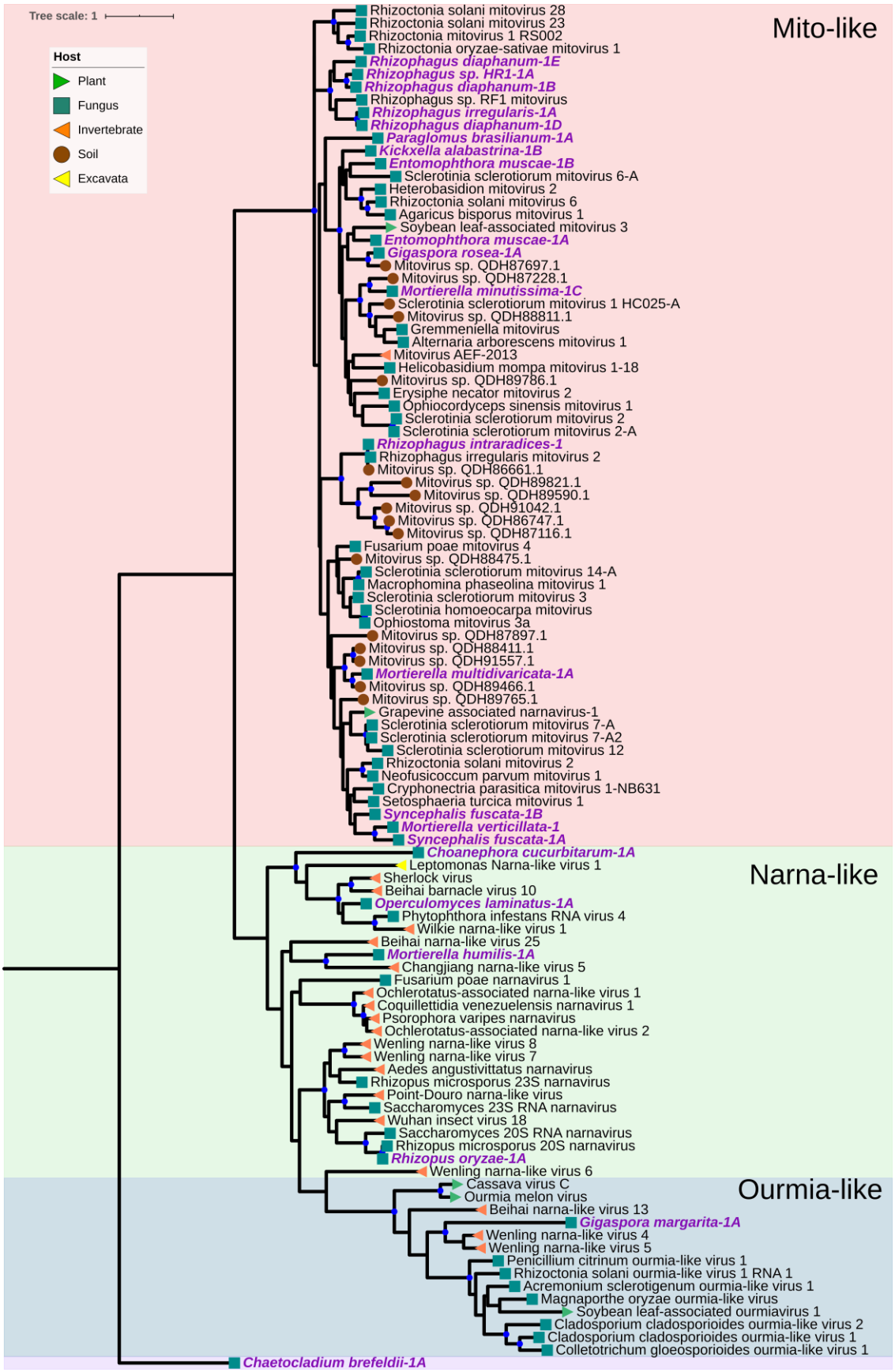

Tree scale: 1
Host
Plant
Fungus
Invertebrate
Soil
Excavata
Mito-like
Narna-like
Ourmia-like
Rhizoctonia solani mitovirus 28
Rhizoctonia solani mitovirus 23
Rhizoctonia mitovirus 1 RS002
Rhizoctonia oryzae-sativae mitovirus 1
Rhizophagus diaphanum-1E
Rhizophagus sp. HR1-1A
Rhizophagus diaphanum-1B
Rhizophagus sp. RF1 mitovirus
Rhizophagus irregularis-1A
Rhizophagus diaphanum-1D
Paraglomus brasilianum-1A
Kickxella alabastrina-1B
Entomophthora muscae-1B
Sclerotinia sclerotiorum mitovirus 6-A
Heterobasidion mitovirus 2
Rhizoctonia solani mitovirus 6
Agaricus bisporus mitovirus 1
Soybean leaf-associated mitovirus 3
Entomophthora muscae-1A
Gigaspora rosea-1A
Mitovirus sp. QDH87697.1
Mitovirus sp. QDH87228.1
Mortierella minutissima-1C
Sclerotinia sclerotiorum mitovirus 1 HC025-A
Mitovirus sp. QDH88811.1
Gremmeniella mitovirus
Alternaria arborescens mitovirus 1
Mitovirus AEF-2013
Helicobasidium mompa mitovirus 1-18
Mitovirus sp. QDH89786.1
Erysiphe necator mitovirus 2
Ophiocordyceps sinensis mitovirus 1
Sclerotinia sclerotiorum mitovirus 2
Sclerotinia sclerotiorum mitovirus 2-A
Rhizophagus intraradices-1
Rhizophagus irregularis mitovirus 2
Mitovirus sp. QDH86661.1
Mitovirus sp. QDH89821.1
Mitovirus sp. QDH89590.1
Mitovirus sp. QDH91042.1
Mitovirus sp. QDH86747.1
Mitovirus sp. QDH87116.1
Fusarium poae mitovirus 4
Mitovirus sp. QDH88475.1
Sclerotinia sclerotiorum mitovirus 14-A
Macrophomina phaseolina mitovirus 1
Sclerotinia sclerotiorum mitovirus 3
Sclerotinia homoeocarpa mitovirus
Ophiostoma mitovirus 3a
Mitovirus sp. QDH87897.1
Mitovirus sp. QDH88411.1
Mitovirus sp. QDH91557.1
Mortierella multidivaricata-1A
Mitovirus sp. QDH89466.1
Mitovirus sp. QDH89765.1
Grapevine associated narnavirus-1
Sclerotinia sclerotiorum mitovirus 7-A
Sclerotinia sclerotiorum mitovirus 7-A2
Sclerotinia sclerotiorum mitovirus 12
Rhizoctonia solani mitovirus 2
Neofusicoccum parvum mitovirus 1
Cryphonectria parasitica mitovirus 1-NB631
Setosphaeria turcica mitovirus 1
Syncephalis fuscata-1B
Mortierella verticillata-1
Syncephalis fuscata-1A
Choanephora cucurbitarum-1A
Leptomonas Narna-like virus 1
Sherlock virus
Beihai barnacle virus 10
Operculomyces laminatus-1A
Phytophthora infestans RNA virus 4
Wilkie narna-like virus 1
Beihai narna-like virus 25
Mortierella humilis-1A
Changjiang narna-like virus 5
Fusarium poae narnavirus 1
Ochlerotatus-associated narna-like virus 1
Coquillettidia venezuelensis narnavirus 1
Psorophora varipes narnavirus
Ochlerotatus-associated narna-like virus 2
Wenling narna-like virus 8
Wenling narna-like virus 7
Aedes angustivittatus narnavirus
Rhizopus microsporus 23S narnavirus
Point-Douro narna-like virus
Saccharomyces 23S RNA narnavirus
Wuhan insect virus 18
Saccharomyces 20S RNA narnavirus
Rhizopus microsporus 20S narnavirus
Rhizopus oryzae-1A
Wenling narna-like virus 6
Cassava virus C
Ourmia melon virus
Beihai narna-like virus 13
Gigaspora margarita-1A
Wenling narna-like virus 4
Wenling narna-like virus 5
Penicillium citrinum ourmia-like virus 1
Rhizoctonia solani ourmia-like virus 1 RNA 1
Acremonium sclerotigenum ourmia-like virus 1
Magnaporthe oryzae ourmia-like virus
Soybean leaf-associated ourmiavirus 1
Cladosporium cladosporioides ourmia-like virus 2
Cladosporium cladosporioides ourmia-like virus 1
Colletotrichum gloeosporioides ourmia-like virus 1
Chaetocladium brefeldii-1A

Figure 2.4 Maximum likelihood tree of new mycovirus RdRps with top blast hits (included in tree) to viruses in "branch 1" of Riboviria. The best model of amino acid substitution for this model was determined to be LG+G per Prottest v. 3.4. Host taxonomy is indicated by branch symbols, and viral taxonomic groupings are indicated by shaded background. Solid blue circles indicate well-supported nodes with ≥ 70% bootstrap support. New sequences are indicated by purple tip labels.

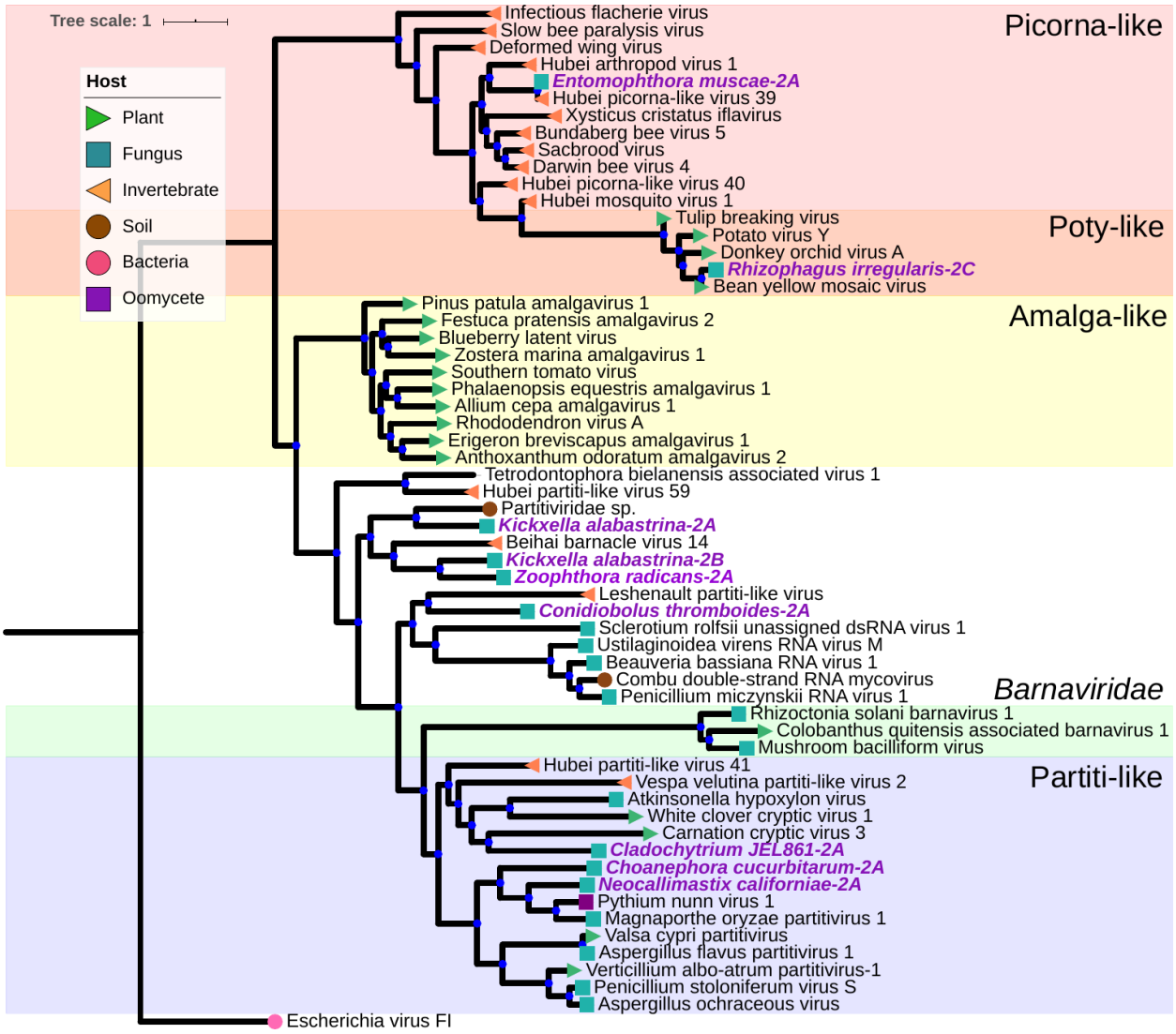


Figure 2.5 Maximum likelihood tree of new mycovirus RdRps with top blast hits (included in tree) to viruses in "branch 2" of Riboviria. The best model of amino acid substitution for this model was determined to be VT+I+G per Prottest v. 3.4. Host taxonomy is indicated by branch symbols, and viral taxonomic groupings are indicated by shaded background. Solid blue circles indicate well-supported nodes with > 70% bootstrap support. New sequences are indicated by purple tip labels.

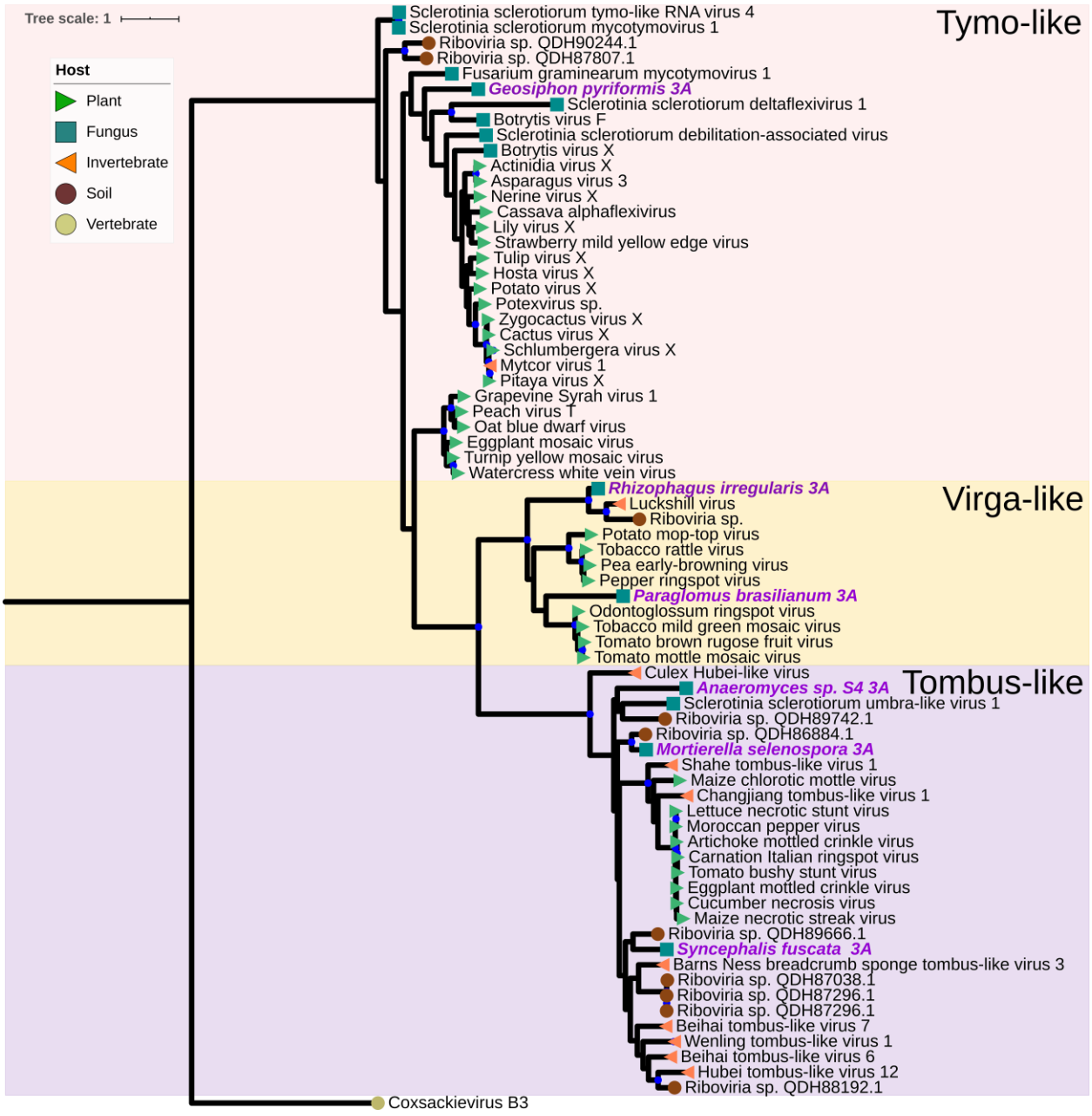


Figure 2.6 Maximum likelihood tree of new mycovirus RdRps with top blast hits (included in tree) to viruses in "branch 3" of Riboviria. The best model of amino acid substitution for this model was determined to be LG+G per Prottest v. 3.4. Host taxonomy is indicated by branch symbols, and viral taxonomic groupings are indicated by shaded background. Solid blue circles indicate well-supported nodes with > 70% bootstrap support. New sequences are indicated by purple tip labels.

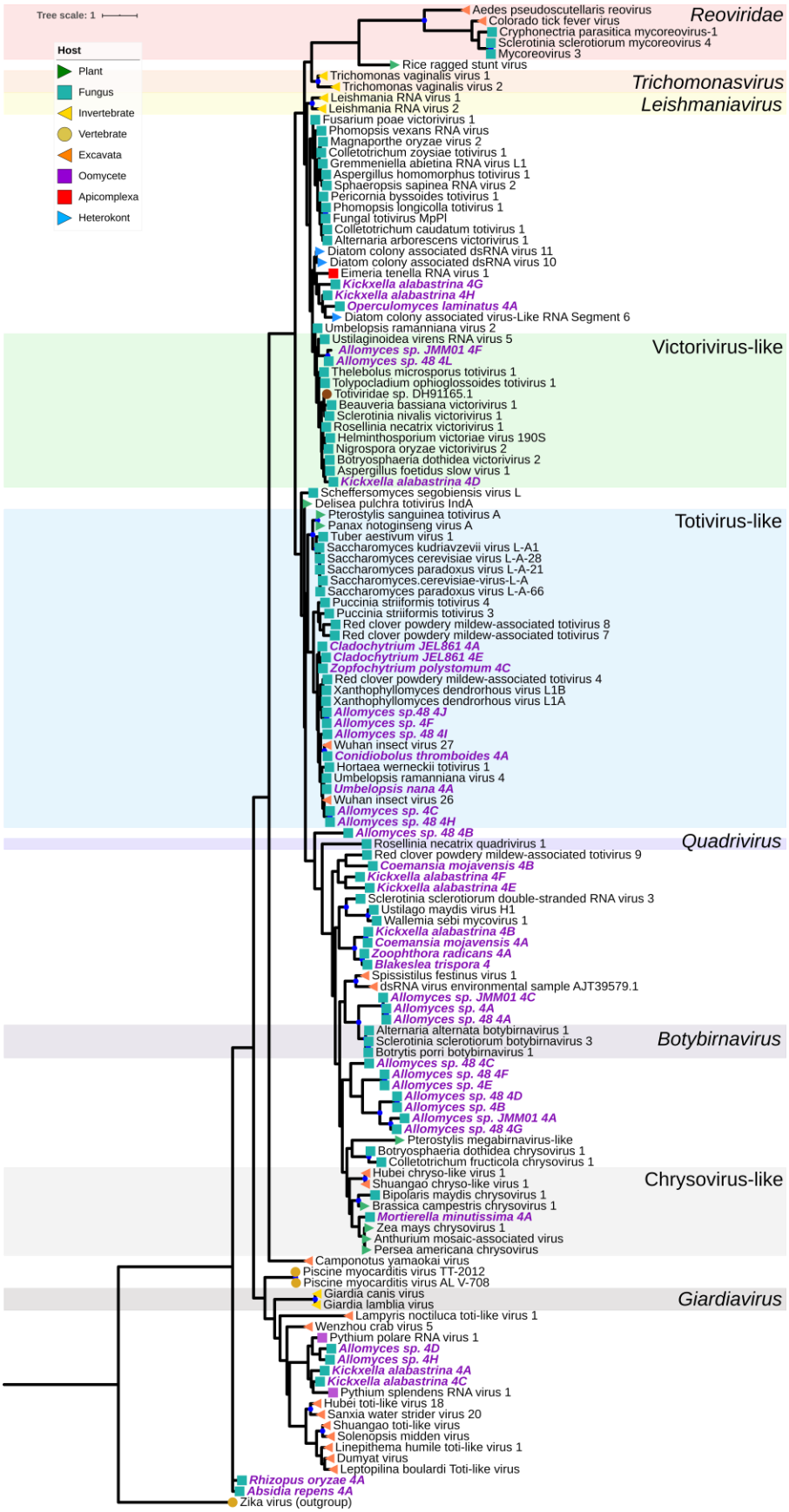

Tree scale: 1
Host
Plant
Fungus
Invertebrate
Vertebrate
Excavata
Oomycete
Apicomplexa
Heterokont
Aedes pseudoscutellaris reovirus
Colorado tick fever virus
Cryphonectria parasitica mycoreovirus-1
Sclerotinia sclerotiorum mycoreovirus 4
Mycoreovirus 3
Reoviridae
Rice ragged stunt virus
Trichomonas vaginalis virus 1
Trichomonas vaginalis virus 2
Trichomonasvirus
Leishmania RNA virus 1
Leishmania RNA virus 2
Leishmaniavirus
Fusarium poae victorivirus 1
Phomopsis vexans RNA virus
Magnaporthe oryzae virus 2
Colletotrichum zoysiae totivirus 1
Gremmeniella abietina RNA virus L1
Aspergillus homomorphus totivirus 1
Sphaeropsis sapinea RNA virus 2
Pericornia byssoides totivirus 1
Phomopsis longicolla totivirus 1
Fungal totivirus MpPI
Colletotrichum caudatum totivirus 1
Alternaria arborescens victorivirus 1
Diatom colony associated dsRNA virus 11
Diatom colony associated dsRNA virus 10
Eimeria tenella RNA virus 1
Kickxella alabastrina 4G
Kickxella alabastrina 4H
Operculomyces laminatus 4A
Diatom colony associated virus-Like RNA Segment 6
Umbelopsis ramanniana virus 2
Ustilaginoidea virens RNA virus 5
Allomyces sp. JMM01 4F
Allomyces sp. 48 4L
Thelebolus microsporus totivirus 1
Tolypocladium ophioglossoides totivirus 1
Totiviridae sp. DH91165.1
Beauveria bassiana victorivirus 1
Sclerotinia nivalis victorivirus 1
Rosellinia necatrix victorivirus 1
Helminthosporium victoriae virus 190S
Nigrospora oryzae victorivirus 2
Botryosphaeria dothidea victorivirus 2
Aspergillus foetidus slow virus 1
Kickxella alabastrina 4D
Victorivirus-like
Scheffersomyces segobiensis virus L
Delisea pulchra totivirus IndA
Pterostylis sanguinea totivirus A
Panax notoginseng virus A
Tuber aestivum virus 1
Saccharomyces kudriavzevii virus L-A1
Saccharomyces cerevisiae virus L-A-28
Saccharomyces paradoxus virus L-A-21
Saccharomyces.cerevisiae-virus-L-A
Saccharomyces paradoxus virus L-A-66
Puccinia striiformis totivirus 4
Puccinia striiformis totivirus 3
Red clover powdery mildew-associated totivirus 8
Red clover powdery mildew-associated totivirus 7
Cladochytrium JEL861 4A
Cladochytrium JEL861 4E
Zopfochytrium polystomum 4C
Red clover powdery mildew-associated totivirus 4
Xanthophyllomyces dendrorhous virus L1B
Xanthophyllomyces dendrorhous virus L1A
Allomyces sp.48 4J
Allomyces sp. 4F
Allomyces sp. 48 4I
Wuhan insect virus 27
Conidiobolus thromboides 4A
Hortaea werneckii totivirus 1
Umbelopsis ramanniana virus 4
Umbelopsis nana 4A
Wuhan insect virus 26
Allomyces sp. 4C
Allomyces sp. 48 4H
Totivirus-like
Allomyces sp. 48 4B
Rosellinia necatrix quadrivirus 1
Red clover powdery mildew-associated totivirus 9
Coemansia mojavensis 4B
Kickxella alabastrina 4F
Kickxella alabastrina 4E
Sclerotinia sclerotiorum double-stranded RNA virus 3
Ustilago maydis virus H1
Wallemia ichthyophaga mycovirus 1
Kickxella alabastrina 4B
Coemansia mojavensis 4A
Zoophthora radicans 4A
Blakeslea trispora 4
Spissistilus festinus virus 1
dsRNA virus environmental sample AJT39579.1
Allomyces sp. JMM01 4C
Allomyces sp. 4A
Allomyces sp. 48 4A
Quadrivirus
Alternaria alternata botybirnavirus 1
Sclerotinia sclerotiorum botybirnavirus 3
Botrytis porri botybirnavirus 1
Allomyces sp. 48 4C
Allomyces sp. 48 4F
Allomyces sp. 4E
Allomyces sp. 48 4D
Allomyces sp. 4B
Allomyces sp. JMM01 4A
Allomyces sp. 48 4G
Botybirnavirus
Pterostylis megabirnavirus-like
Botryosphaeria dothidea chrysovirus 1
Colletotrichum fructicola chrysovirus 1
Hubei chryso-like virus 1
Shuangao chryso-like virus 1
Bipolaris maydis chrysovirus 1
Brassica campestris chrysovirus 1
Mortierella minutissima 4A
Zea mays chrysovirus 1
Anthurium mosaic-associated virus
Persea americana chrysovirus
Chrysovirus-like
Camponotus yamaokai virus
Piscine myocarditis virus TT-2012
Piscine myocarditis virus AL V-708
Giardia canis virus
Giardia lamblia virus
Giardiavirus
Lampyris noctiluca toti-like virus 1
Wenzhou crab virus 5
Pythium polare RNA virus 1
Allomyces sp. 4D
Allomyces sp. 4H
Kickxella alabastrina 4A
Kickxella alabastrina 4C
Pythium splendens RNA virus 1
Hubei toti-like virus 18
Sanxia water strider virus 20
Shuangao toti-like virus
Solenopsis midden virus
Linepithema humile toti-like virus 1
Dumyat virus
Leptopilina boulardi Toti-like virus
Rhizopus oryzae 4A
Absidia repens 4A
Zika virus (outgroup)

Figure 2.7 Maximum likelihood tree of new mycovirus RdRps with top blast hits (included in tree) to viruses in "branch 4" of Riboviria. The best model of amino acid substitution for this model was determined to be LG+G per Prottest v. 3.4. Host taxonomy is indicated by branch symbols, and viral taxonomic groupings are indicated by shaded background. Solid blue circles indicate well-supported nodes with > 70% bootstrap support. New sequences are indicated by purple tip labels.

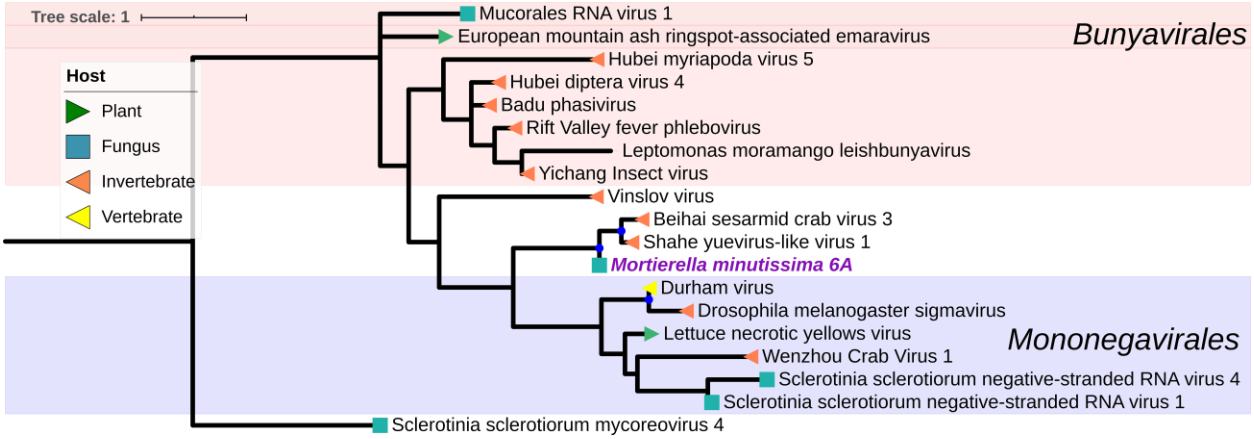


Figure 2.8 Maximum likelihood tree of a new mycovirus RdRp with top blast hits (included in tree) to viruses in "branch 5" of Riboviria. The best model of amino acid substitution for this model was determined to be VT+G per Prottest v. 3.4. Host taxonomy is indicated by branch symbols, and viral taxonomic groupings are indicated by shaded background. Solid blue circles indicate well-supported nodes with > 70% bootstrap support. New sequences are indicated by purple tip labels.

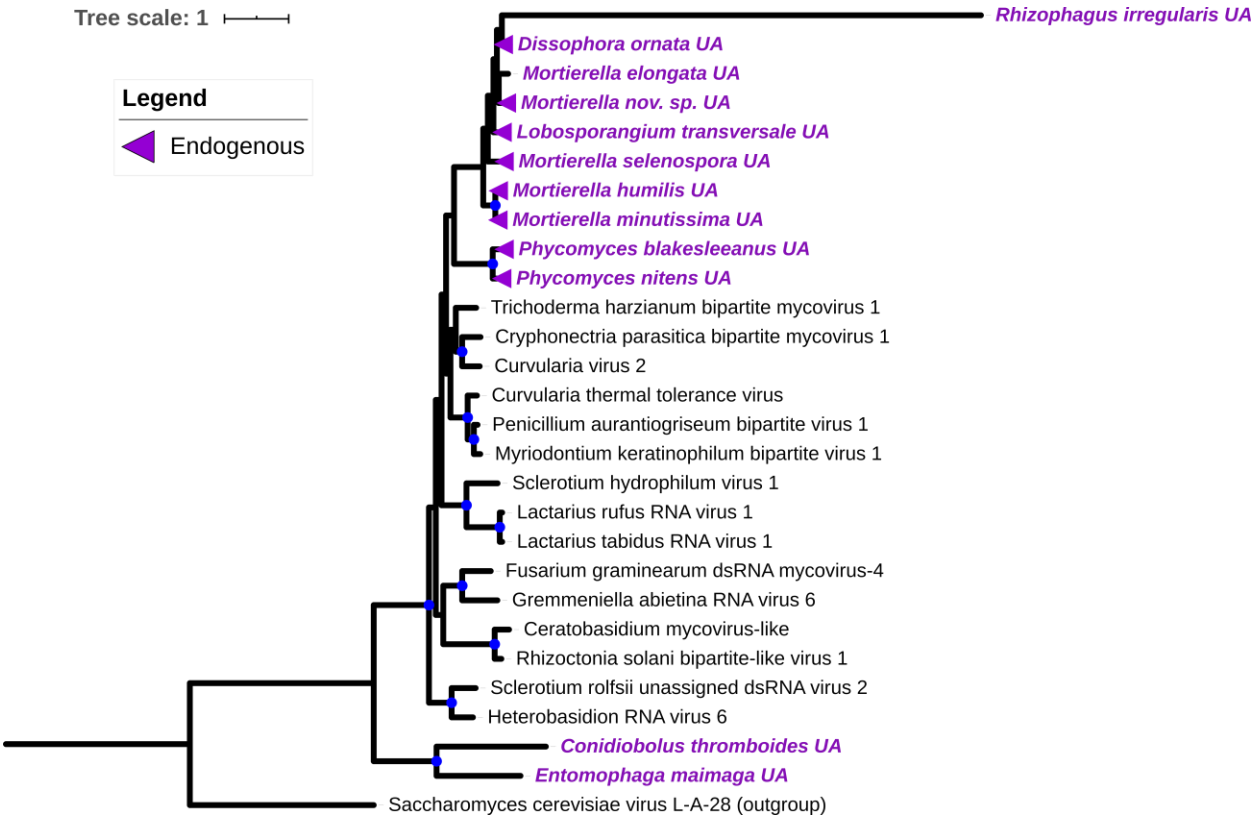


Figure 2.9 Maximum-likelihood tree of new mycovirus RdRps with top blast hits (included in tree) unassigned by current viral taxonomy. The best model of amino acid substitution for this model was determined to be LG +G per Prottest v. 3.4. All viruses have fungal hosts. New sequences are indicated by purple tip labels. Blue circles indicate nodes with bootstrap support > 70%. Triangles indicate novel viral-like sequences determined to be endogenous in the host genome. A DNA-based genome for *Mortierella elongata* has yet to be sequenced, and so we cannot conclude that *Mortierella elongata UA* virus is endogenous, although it appears likely.

Tables

Table 2.1 Percent viral prevalence based on combined *in vitro* and *in silico* screening, by phyla and sub-phyla, or order.

Taxonomic group	Sample size (n)	Percent virus prevalence (%)
Mucoromycota	107	27.1
Glomeromycotina	9	88.9
Mucoromycotina	71	19.7
Mortierellomycotina	27	25.9
Zoopagomycota	25	40.0
Entomophthoromycotina	11	54.5
Kickxellomycotina	11	27.3
Zoopagomycotina	3	33.3
Blastocladiomycota	44	15.9
Blastocladiales	42	16.7
Physodermatales	2	0
Chytridiomycota	146	16.4
Monoblepharidales	5	0
Cladochytriales	21	47.6
Synchytriales	2	0
Chytridiales	40	10.0
Lobulomycetales	2	0
Rhizophydiales	35	14.3
Rhizophlyctidales	6	0
Spizellomycetales	36	13.9
Neocallimastigomycota	8	25.0
Cryptomycota	2	0
Total	333	21.6

References

Ghabrial, S. A., Castón, J. R., Jiang, D., Nibert, M. L., & Suzuki, N. (2015). 50-Plus Years of Fungal Viruses. *Virology*, *479–480*, 356–368. https://doi.org/10.1016/j.virol.2015.02.034

Son, M., Yu, J., & Kim, K. H. (2015). Five Questions about Mycoviruses. *PLoS Pathogens*, *11*(11), 5–11. https://doi.org/10.1371/journal.ppat.1005172

James, T. Y., J. E. Stajich, C. T. Hittinger, and A. Rokas. (2020). Towards a fully resolved fungal tree of life. *Annu. Rev. Microbiol*. 74: 291-313.

Spatafora, J., Aime, M., Grigoriev, I., Martin, F., Stajich, J., & Blackwell, M. (2017). The fungal tree of life: from molecular systematics to genome-scale phylogenies. *Microbiology Spectrum* 5(5): doi:10.1128/microbiolspec.FUNK-0053-2016

Hibbett, D. S., Binder, M., Bischoff, J. F., Blackwell, M., Cannon, P. F., Eriksson, O. E., … Zhang, N. (2007). A higher-level phylogenetic classification of the Fungi. *Mycological Research*. https://doi.org/10.1016/j.mycres.2007.03.004

Khandjian, E.W., Roos, U., Timberlake, W.E., Eder, L., & Turian, G. (1974). RNA Virus-Like Particles in the Chytriomycete *Allomyces arbuscula. Archives of Microbiology, 101*. 351-356.

Papp, T., Nyilasi, I., Fekete, C., Ferenczy, L., & Vágvölgyi, C. (2001). Presence of double-stranded RNA and virus-like particles in *Rhizopus* isolates. *Candian Journal of Microbiology, 47*(5), 443–447.

Kitahara, R., Ikeda, Y., Shimura, H., Masuta, C., & Ezawa, T. (2014). A unique mitovirus from Glomeromycota, the phylum of arbuscular mycorrhizal fungi. *Archives of Virology*, *159*(8), 2157–2160. https://doi.org/10.1007/s00705-014-1999-1

Neupane, A., Feng C., Feng J., Kafle A., Bücking, H., and Marzano, S.L. (2018). Metatranscriptomic analysis and in silico approach identified mycoviruses in the arbuscular mycorrhizal fungus *Rhizophagus* Spp. *Viruses 10* (12). https://doi.org/10.3390/v10120707.

Kartali, T., Nyilasi, I., Szabó, B., Kocsubé, S., Patai, R., Polgár, T. F., … Papp, T. (2019). Detection and Molecular Characterization of Novel dsRNA Viruses Related to the Totiviridae Family in *Umbelopsis ramanniana*. *Frontiers in Cellular and Infection Microbiology*, *9*(July), 1–13. https://doi.org/10.3389/fcimb.2019.00249

Coyle, M. C., Elya, C. N., Bronski, M., & Eisen, M. B. (2018). Entomophthovirus: An insect-derived iflavirus that infects a behavior manipulating fungal pathogen of dipterans. *BioRxiv*, 371526. https://doi.org/10.1101/371526

Nibert, M. L., Debat, H. J., Manny, A. R., Grigoriev, I. V., & De Fine Licht, H. H. (2019). Mitovirus and mitochondrial coding sequences from basal fungus *Entomophthora muscae*. *Viruses*, *11*(4), 6–8. https://doi.org/10.3390/v11040351

Espino-Vázquez, A.N., Bermúdez-Barrientos, J.R., Cabrera-Rangel, J.F.*et al*.(2020).

Narnaviruses: novel players in fungal–bacterial symbioses. *ISME J 14***,** 1743–1754. https://doi.org/10.1038/s41396-020-0638-y

Morris, T. J., & Dodds, J. A. (1979). Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. *Phytopathology*, 69(8), 854–858.

Nerva, L. *et al.* (2015). Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi. *Virus Res. 219*, 22–38.

Marzano, S. Y. L. & Domier, L. L. (2016) Novel mycoviruses discovered from metatranscriptomics survey of soybean phyllosphere phytobiomes. *Virus Res. 219*, 11–21.

Gilbert, K. B., Holcomb, E. E., Allscheid, R. L. & Carrington, J. C. (2019). Hiding in plain sight: New virus genomes discovered via a systematic analysis of fungal public transcriptomes. *Plos One 14.*

Osaki, H., Sasaki, A., Nomiyama, K. & Tomioka, K. (2016). Multiple virus infection in a single strain of *Fusarium poae* shown by deep sequencing. *Virus Genes 52*, 835–847.

Okada, R., Kiyota, E., Moriyama, H., Fukuhara, T., & Natsuaki, T. 2015. A simple and rapid method to purify viral dsRNA from plant and fungal tissue. *Journal of General Plant Pathology 81* (2): 103–7. https://doi.org/10.1007/s10327-014-0575-6.

Hannon, G.J. (2010). FASTX-Toolkit. http://hannonlab.cshl.edu/fastx_toolkit.

Grabherr M.G., Haas B.J., Yassour M., Levin J.Z., Thompson D.A., Amit I., *et al.* (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology, 29***:**7, 644–652. doi: 10.1038/nbt.1883c.

Haas, B.J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P.D., Bowden, J., *et al.* (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, **8** (8). 1494–512. doi:10.1038/nprot.2013.084.

Wheeler, T.S., & Eddy, S. R. (2013). nhmmer: DNA homology search with profile HMMs. *Bioinformatics*, *29* (19). 2487–2489. https://doi.org/10.1093/bioinformatics/btt403.

Ruby, J. G., Bellare, P. & DeRisi, J. L. (2013). PRICE: Software for the targeted assembly of components of (meta) genomic sequence data. *G3 Genes, Genomes, Genet. 3*, 865–880.

MacManes, M. D. (2014). On the optimal trimming of high-throughput mRNA sequence data. *Front. Genet. 5*, 1–7.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15–21. https://doi.org/10.1093/bioinformatics/bts635.

Sasai, S. *et al.* (2018). A novel non-segmented double-stranded RNA virus from an Arctic isolate

of Pythium polare. *Virology 522*, 234–243.

Wolf, Y. I. *et al.* (2018). Origins and evolution of the global RNA virome. *MBio 9*, 1–31.

Katoh, K, & DM Standley. (2013). MAFFT: multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution 30* (4): 772–80. https://doi.org/10.1093/molbev/mst010.

Capella-Gutiérrez, S., Silla-Martínez, J. M. & Gabaldón, T. (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics 25*, 1972–1973.

Darriba, D, GL Taboada, R Doallo, and D Posada. (2017). ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics 27* (8): 1164–65. https://doi.org/10.1093/bioinformatics/btr088.ProtTest.

Stamatakis, A. (2014). RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics 30* (9): 1312–13. https://doi.org/10.1093/bioinformatics/btu033.

Sutela, S., Poimala, A. & Vainio, E. J. (2019). Viruses of fungi and oomycetes in the soil environment. *FEMS Microbiol. Ecol. 95*, 1–18.

Bonfante, P, and A Desirò. (2017). Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. *ISME Journal 11* (8): 1727–35. https://doi.org/10.1038/ismej.2017.21.

Bozarth, R. F. (1972). Mycoviruses : A New Dimension In Microbiology. 23–39.

Nibert, M. L., Vong, M., Fugate, K. K. & Debat, H. J. (2018). Evidence for contemporary plant mitoviruses. *Virology 518*, 14–24.

Göker, M, C Scheuner, HP Klenk, JB Stielow, and W Menzel. (2011). Codivergence of mycoviruses with their hosts. *PLoS ONE 6* https://doi.org/10.1371/journal.pone.0022252.

Márquez, LM, RS Redman, RJ Rodriguez, and MJ Roossinck. (2007). A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance. *Science 315* (5811): 513–15.

Frank, A. C. & Wolfe, K. H. (2009). Evolutionary capture of viral and plasmid DNA by yeast nuclear Chromosomes. *Eukaryot. Cell 8*, 1521–1531.

Taylor, D. J. & Bruenn, J. (2009). The evolution of novel fungal genes from non-retroviral RNA viruses. *BMC Biol. 7*, 88.

Liu, H. *et al.* (2010). Widespread horizontal gene transfer from double-stranded RNA viruses to eukaryotic nuclear genomes. *J. Virol. 84*, 11876–87

Supplement

Supplemental Table 2.1 Viral screening results for all isolates screened by both culture-based and transcriptome-mining approaches. Asterisks denote viral positives determined to be endogenous in the host genome.

Phylum	Taxon	Isolate	Screening method	Virus pos/neg	Num. bands	Num. Rdrp	Taxa Rdrp
Mucoromycota	Mucoromycotina	*Absidia californica* NRRL 2967	in vivo	neg			
Mucoromycota	Mucoromycotina	*Absidia corymbifera* NRRL A-14836	in vivo	neg			
Mucoromycota	Mucoromycotina	*Absidia ramosa* NRRL 3180	in vivo	neg			
Mucoromycota	Mucoromycotina	*Absidia repens* NRRL 1336	in silico	pos		1	Branch 4
Mucoromycota	Mucoromycotina	*Actinomucor elegans* NRRL 3104	in vivo	pos	1		
Blastocladiomycota	Blastocladiales	*Allomyces* sp. DJ02	in vivo, in silico	pos, pos	6	6	Branch 4
Blastocladiomycota	Blastocladiales	*Allomyces* sp. DJ05	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces* sp. DJ07	in vivo, in silico	pos, pos	4	4	Branch 4
Blastocladiomycota	Blastocladiales	*Allomyces* sp. WWM105	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Brit East Africa 2	in vivo	pos	1		
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* ATCC 10983	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Costa Rica 18	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Cuba S 20	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Mexico 57	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Ohio 5	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Cuba S 22	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* FGSC48	in vivo	pos	8		
Blastocladiomycota	Blastocladiales	*Allomyces javanicus* Cuba S 12	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces macrogynus* ATCC_38327	in silico	neg			
Blastocladiomycota	Blastocladiales	*Allomyces moniliformis* JMM01	in vivo, in silico	pos, pos	>5	8	Branch 4

Blastocladiomycota	Blastocladiales	*Allomyces javanicus* India-4	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces macrogynus* Australia-3	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces anomalus* California-6	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces* sp. BEA	in vivo	pos	2		
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Burma 1F	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Belgian Congo 1	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Burma 1Db	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* California-7	in vivo	pos	2		
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* CBS10.463	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Costa Rica 57A	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Cuba S 28	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Denmark 1	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* El Salvador 1	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Florida F2/F12	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* North Carolina 2	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces arbuscula* Philippines Isl 1	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces javanicus* California-1	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces* sp. Australia CaC	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Allomyces* sp. Fla 2	in vivo	neg			
Mucoromycota	Mucoromycotina	*Amylomyces rouxii* NRRL 5866	in silico	neg			
Mucoromycota	Mucoromycotina	*Amylomyces rouxii* NRRL A-11375	in vivo	neg			
Mucoromycota	Mucoromycotina	*Amylomyces rouxii* NRRL A-25885	in vivo	neg			
Mucoromycota	Mucoromycotina	*Amylomyces rouxii* NRRL A-26221	in vivo	neg			
Neocallimastigomycota	Neocallimastigales	*Anaeromyces* sp. NHY-2018	in silico	neg			
Neocallimastigomycota	Neocallimastigales	*Anaeromyces* sp. S4	in silico	pos		1	Branch 3
Mucoromycota	Mucoromycotina	*Backusella circina* 2446	in vivo	neg			

Mucoromycota	Mucoromycotina	*Backusella ctenidia* 6239	in vivo	neg			
Zoopagomycota	Entomophthoromycotina	*Basidiobolus* sp. JELUM001	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* 423	in silico	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* Campana13	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* Campana24	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* CLFT044	in silico	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* FMB03	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* NAF077	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* NMBF05	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* SAF014	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium dendrobatidis* UM142	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Batrachochytrium salamandrivorans*	in silico	neg			
Mucoromycota	Mucoromycotina	*Benjaminiella poitrasii* RSA 903	in silico	neg			
Mucoromycota	Mucoromycotina	*Bifiguratus adelaidae* AZ 501	in vivo	neg			
Mucoromycota	Mucoromycotina	*Bifiguratus adelaidae* TLT 265	in vivo	neg			
Mucoromycota	Mucoromycotina	*Blakeslea trispora* F986	in silico	pos	2	Branch 4	
Mucoromycota	Mucoromycotina	*Blakeslea trispora* NRRL 2456	in vivo, in silico	pos, neg	1		
Mucoromycota	Mucoromycotina	*Blakeslea trispora* NRRL 2895	in vivo	pos	1		
Blastocladiomycota	Blastocladiales	*Blastocladiella britanica* JEL0711	in vivo, in silico	neg, neg			
Blastocladiomycota	Blastocladiales	*Blastocladiella brittanica* Barr214	in vivo	neg			
Chytridiomycota	Chytridiales	*Blyttiomyces* sp. JEL0837	in vivo	neg			
Chytridiomycota	Rhizophlyctidales	*Boealophlyctis nickersoniae* WJD171	in vivo	neg			

Chytridiomycota	Rhizophlyctidales	*Borealophlyctis nickersonii* WJD170	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Catenaria* sp. JEL0871	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Catenaria* sp. JEL748	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Catenaria* sp. MP54	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Catenaria* sp. MP55	in vivo	neg			
Blastocladiomycota	Blastocladiales	*Catenaria* sp. PL171	in vivo, in silico	neg, neg			
Chytridiomycota	Chytridiales	*Catenochytridium* sp. JEL0775	in vivo	neg			
Mucoromycota	Mucoromycotina	*Chaetocladium brefeldii* NRRL 2508	in vivo	neg			
Mucoromycota	Mucoromycotina	*Chaetocladium brefeldii* NRRL 2343	in silico	pos		1	Branch 1
Mucoromycota	Mucoromycotina	*Chlamydoabsidia padenii* NRRL 2977	in silico	neg			
Mucoromycota	Mucoromycotina	*Choanephora cucurbitarum* NRRL 2744	in silico	pos		1	Branch 2
Chytridiomycota	Chytridiales	*Chytridium lagenaria* Arg66	in vivo, in silico	neg, neg			
Chytridiomycota	Chytridiales	*Chytriomyces hyalinus* JEL0117	in vivo	neg			
Chytridiomycota	Chytridiales	*Chytriomyces hyalinus* JEL0632	in vivo, in silico	neg, neg			
Chytridiomycota	Chytridiales	*Chytriomyces* sp. nov. MP71	in silico	neg			
Mucoromycota	Mucoromycotina	*Circinella angarensis* NRRL 2628	in vivo	pos	1		
Mucoromycota	Mucoromycotina	*Circinella minor* NRRL A-13969	in vivo	neg			
Mucoromycota	Mucoromycotina	*Circinella muscae* NRRL 1364	in vivo	neg			
Mucoromycota	Mucoromycotina	*Circinella naumovii* NRRL 5846	in vivo	neg			
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0144	in vivo	pos	2		
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0153	in vivo	neg			
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0477	in vivo	pos	1		
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0479	in vivo	neg			
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0521	in vivo	pos	5		
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0592	in vivo	pos	1		

Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0704	in vivo	pos	1		
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0714	in vivo	pos	5		
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0772	in vivo	pos	2		
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0861	in vivo	pos	>8		
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0893	in vivo	neg			
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0899	in vivo	neg			
Chytridiomycota	Cladochytriales	*Cladochytrium* sp JEL0900	in vivo	neg			
Chytridiomycota	Cladochytriales	*Cladochytrium* sp. JEL0903	in vivo	neg			
Chytridiomycota	Cladochytriales	*Cladochytrium tenue* GHJ CCIBt 4013	in vivo	neg			
Mucoromycota	Glomeromycotina	*Claroideoglomus etunicatum*	in silico	neg			
Blastocladiomycota	Blastocladiales	*Coelomomyces lativattus* CIRM-AVA-1	in silico	neg			
Zoopagomycota	Kickxellomyocotina	*Coemansia mojavensis* RSA 71	in silico	pos		2	Branch 4
Zoopagomycota	Kickxellomyocotina	*Coemansia reversa* NRRL1564	in vivo	pos	3		
Zoopagomycota	Kickxellomyocotina	*Coemansia spiralis* RSA 1278	in vivo, in silico	neg, neg			
Mucoromycota	Mucoromycotina	*Cokeromyces recurvatus* NRRL 2243	in silico	neg			
Zoopagomycota	Entomophthoromycotina	*Conidiobolus antarcticus* ARSEF 6913	in vivo	neg			
Zoopagomycota	Entomophthoromycotina	*Conidiobolus coronatus* ARSEF 9914	in vivo	pos	1		
Zoopagomycota	Entomophthoromycotina	*Conidiobolus coronatus* NRRL 28638	in silico	neg			
Zoopagomycota	Entomophthoromycotina	*Conidiobolus coronatus* WDUM101	in vivo	neg			
Zoopagomycota	Entomophthoromycotina	*Conidiobolus coronatus* WDUM102	in vivo	neg			
Zoopagomycota	Entomophthoromycotina	*Conidiobolus thromboides* FSU 785	in silico	pos		5	Branch 4 (4), Branch 2 (1)
Mucoromycota	Mucoromycotina	*Cunninghamella* sp. NRRL A-21271	in vivo	neg			
Mucoromycota	Mucoromycotina	*Dichotomocladium elegans* RSA 919	in silico	neg			

Mucoromycota	Mortierellomycotina	*Dissophora decumbens* NRRL 22416	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Dissophora ornata* CBS 347.77	in silico	pos		1	Unassigned
Chytridiomycota	Chytridiales	*Endochytrium* sp. JEL0050	in vivo	neg			
Chytridiomycota	Chytridiales	*Endochytrium* sp. JEL0386	in vivo	neg			
Chytridiomycota	Chytridiales	*Endochytrium* sp. JEL0896	in vivo	neg			
Zoopagomycota	Entomophthoromycotina	*Entomophaga maimaga* ARSEF 7190	in silico	pos		1	Unassigned
Zoopagomycota	Entomophthoromycotina	*Entomophthora muscae* HHdFL130914-01	in silico	pos		4	Branch1 (3), Branch 2 (1)
Chytridiomycota	Chytridiales	*Entophlyctis helioformis* JEL0805	in vivo, in silico	neg, neg			
Chytridiomycota	Chytridiales	*Entophlyctis* sp. JEL0112	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Fimicolochytrium alabamae* JEL0588	in vivo	pos	2		
Chytridiomycota	Spizellomycetes	*Fimicolochytrium* sp. JEL0733	in vivo	pos	1		
Zoopagomycota	Kickxellomyocotina	*Furculomyces boomerangus* NRRL 9021	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces* sp. JEL0119	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces* sp. JEL0628	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces* sp. JEL0657	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces* sp. JEL0662	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces* sp. JEL0698	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces* sp. JEL0699	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces* sp. JEL0738	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Gaertneriomyces semiglobifer* Barr 43	in vivo	neg			
Mucoromycota	Glomeromycotina	*Geosiphon pyriformis*	in silico	pos		1	Branch 3
Chytridiomycota	Spizellomycetes	*Geranomyces michiganensis* JEL0563	in vivo	pos	2		
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* Barr350	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* JEL 559	in silico	neg			
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* JEL0542	in vivo	neg			

Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* JEL0557	in vivo	pos	1		
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* JEL0566	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* JEL0567	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* KP27	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* KP31	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* MP0003	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Geranomyces variabilis* MP0004	in vivo	neg			
Mucoromycota	Glomeromycotina	*Gigaspora margarita* BEG34	in silico	pos		1	Branch 1
Mucoromycota	Glomeromycotina	*Gigaspora rosea*	in silico	pos		3	Branch 1
Mucoromycota	Mucoromycotina	*Gilbertella persicaria* NRRL A-13613	in vivo	pos	1		
Mucoromycota	Mucoromycotina	*Gilbertella persicaria* var. *persicaria* CBS 190.32	in silico	neg			
Chytridiomycota	Rhizophydiales	*Globomyces pollinis-pini* Arg68	in vivo, in silico	neg, neg			
Chytridiomycota	Monoblepharidales	*Gonapodya* sp. JEL0183	in vivo	neg			
Mucoromycota	Mucoromycotina	*Gongronella butleri* NRRL 1340	in vivo	neg			
Mucoromycota	Mucoromycotina	*Gongronella butleri* NRRL A-23795	in vivo	neg			
Mucoromycota	Mucoromycotina	*Gongronella butleri* C1D	in silico	neg			
Chytridiomycota	Rhizophydiales	*Gorgonomyces* sp. Arg29	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Gorgonomyces haynaldii* MP0057	in vivo, in silico	neg, neg			
Mucoromycota	Mucoromycotina	*Halteromyces radiatus* CBS 162.75 T	in silico	neg			
Chytridiomycota	Monoblepharidales	*Harpochytrium* sp. JEL0705	in vivo	neg			
Mucoromycota	Mucoromycotina	*Helicostylum pulchrum* RSA 2064	in silico	neg			
Mucoromycota	Mucoromycotina	*Hesseltinella vesiculosa* NRRL 3301	in silico	neg			
Mucoromycota	Mucoromycotina	*Hesseltinella vesiculosa* NRRL 34	in vivo	neg			
Chytridiomycota	Monoblepharidales	*Hyaloraphidium curvatum* JEL0383	in vivo, in silico	neg, neg			
Chytridiomycota	Rhizophydiales	*Kappamyces* sp. JEL0680	in vivo	pos			

Zoopagomycota	Kickxellomyocotina	*Kickxella alabastrina* RSA 675	in silico	pos		15	Branch 4 (9), Branch 2 (4), Branch 1 (2)
Mucoromycota	Mucoromycotina	*Kirkomyces cordense* RSA 1222	in silico	neg			
Zoopagomycota	Kickxellomyocotina	*Legeriosimilis* sp. ARSEF 9066	in vivo	neg			
Zoopagomycota	Kickxellomyocotina	*Linderina pennispora* ATCC12442/NRRL2237	in vivo, in silico	neg, neg			
Mucoromycota	Mortierellomycotina	*Lobosporangium transversale* NRRL 3116	in silico	pos		1	Unassigned
Chytridiomycota	Lobulomycetales	*Lobulomyces angularis* JEL0522	in vivo	neg			
Chytridiomycota	Lobulomycetales	*Lobulomyces poculatus* JEL0511	in vivo	neg			
Zoopagomycota	Kickxellomyocotina	*Martensiomyces pterosporus* CBS 209.56	in silico	neg			
Mucoromycota	Mortierellomycotina	*Mortierella alpina*	in silico	neg			
Mucoromycota	Mortierellomycotina	*Mortierella ambigua* NRRL 28271	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella capitata* AV 005	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella echinospaera* NRRL 1233	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella elongata* NRRL 5513	in silico	pos		1	Unassigned
Mucoromycota	Mortierellomycotina	*Mortierella elongata* NVP 64	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella epicladia* AD 058	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella gamsii* AM 1032	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella humilis* PMI 1414	in vivo, in silico	neg, neg			
Mucoromycota	Mortierellomycotina	*Mortierella hyalina* NRRL 6427	in vivo, in silico	neg, pos		2	Branch 1, Unassigned
Mucoromycota	Mortierellomycotina	*Mortierella minutissima* AD051	in vivo	pos	1		
Mucoromycota	Mortierellomycotina	*Mortierella multidivaricata* NRRL 6456	in silico	pos		6	Branch 1 (4), Branch 4 (1), Unassigned (1)
Mucoromycota	Mortierellomycotina	*Mortierella multidivaricata* RSA 2152	in vivo	neg			

Mucoromycota	Mortierellomycotina	*Mortierella* nov. sp. GBAus 27b	in silico	pos		4	Branch 1
Mucoromycota	Mortierellomycotina	*Mortierella polycephala* KOD 948	in vivo, in silico	neg, pos		1	Unassigned
Mucoromycota	Mortierellomycotina	*Mortierella selenospora* CBS 811.68	in vivo	pos	2		
Mucoromycota	Mortierellomycotina	*Mortierella selenospora* KOD 1015	in silico	pos		2	Branch 3 (1), Unassigned (1)
Mucoromycota	Mortierellomycotina	*Mortierella* sp. GBAus 30	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella* sp. JEL0843	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella* sp. JEL0858	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella* sp. JEL0860	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella verticillata* NRRL 6337	in vivo	neg			
Mucoromycota	Mortierellomycotina	*Mortierella wolfii* NRRL 28640	in silico	pos		2	Branch 1
Mucoromycota	Mucoromycotina	*Mucor circinelloides*	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor griseo-ochraceus* var. *minuta* NRRL 3246	in silico	neg			
Mucoromycota	Mucoromycotina	*Mucor heimalis* WDUM104 777	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor hiemalis f. hiemalis* NRRL 3624	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor irregularis* C3B	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor lahorensis* NRRL 6592	in silico	neg			
Mucoromycota	Mucoromycotina	*Mucor mousanensis* NRRL 3105	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor pakistanicus* NRRL 6589	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor petrinsularius* var. *echinosporus* NRRL 3141	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor petrinsularius* var. *ovalisporus* NRRL 2536	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor ramosissimus* NRRL 3042	in vivo	pos	2		
Mucoromycota	Mucoromycotina	*Mucor recurvus* NRRL 2358	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor rouxii* NRRL 3367	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucor rouxii* NRRL A-11340	in vivo	neg			

Mucoromycota	Mucoromycotina	*Mucor rouxii* NRRL A-11341	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mucorales* sp. UM774	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mycocladus corymbifer* NRRL 1309	in vivo	neg			
Mucoromycota	Mucoromycotina	*Mycotypha africana* NRRL 2978	in vivo	pos	2		
Mucoromycota	Mucoromycotina	*Mycotypha microspora* NRRL 684	in silico	neg			
Neocallimastigomycota	Neocallimastigales	*Neocallimastix californiae*	in vivo	neg			
Neocallimastigomycota	Neocallimastigales	*Neocallimastix frontalis* 27	in silico	pos		1	Branch 2
Chytridiomycota	Cladochytriales	*Nephrochytrium aurantium* JEL0909	in silico	neg			
Chytridiomycota	Cladochytriales	*Nowakowskiella* sp. JEL0774	in vivo	neg			
Chytridiomycota	Cladochytriales	*Nowakowskiella* sp. JEL0785	in vivo	pos	1		
Chytridiomycota	Chytridiales	*Obelidium* sp. JEL0802	in vivo	pos	1		
Chytridiomycota	Rhizophydiales	*Operculomyces laminatus* JEL223	in vivo, in silico	neg, neg			
Neocallimastigomycota	Neocallimastigales	*Orpinomyces* sp. C1A	in silico	pos		2	Branch 1; Branch 4
Neocallimastigomycota	Neocallimastigales	*Orpinomyces joyonii* SG4	in silico	neg			
Mucoromycota	Glomeromycotina	*Paraglomus brasilianum*	in silico	neg			
Cryptomycota		*Paramicrosporidium saccamoebae*	in silico	pos		2	Branch 1, Branch 3
Chytridiomycota	Rhizophydiales	*Paranamyces uniporus* JEL0695	in silico	neg			
Chytridiomycota	Rhizophydiales	*Paranomyces uniporus* WJD150	in vivo	pos	1		
Chytridiomycota	Rhizophydiales	*Paranomyces uniporus* WJD158	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Paranomyces uniporus* WJD193	in vivo	pos	1		
Blastocladiomycota	Physodermatales	*Paraphysoderma sedebokerense* JEL0821	in vivo	neg			
Blastocladiomycota	Physodermatales	*Paraphysoderma sedebokerense* JEL0847	in vivo, in silico	neg, neg			
Mucoromycota	Mucoromycotina	*Parasitella parasitica* NRRL 2501	in vivo	neg			
Mucoromycota	Mucoromycotina	*Phascolomyces articulosus* RSA 2281	in silico	neg			
Mucoromycota	Mucoromycotina	*Phycomyces blakesacanus* UM 175	in silico	neg			

Mucoromycota	Mucoromycotina	*Phycomyces blakesleeanus*	in vivo	neg			
Mucoromycota	Mucoromycotina	*Phycomyces nitens* S609	in silico	pos		1	Unassigned
Chytridiomycota	Chytridiales	*Phylctochytrium bullatum* JEL0754	in silico	pos		1	Unassigned
Chytridiomycota	Chytridiales	*Phylctochytrium planicorne* JEL0894	in vivo	neg			
Chytridiomycota	Chytridiales	*Physocladia obscura* JEL0137	in vivo	neg			
Zoopagomycota	Zoopagomycotina	*Piptocephalis cylindrospora* RSA 2659	in vivo	neg			
Neocallimastigomycota	Neocallimastigales	*Piromyces rhizinflatus* YM600	in silico	neg			
Neocallimastigomycota	Neocallimastigales	*Piromyces finn*	in silico	neg			
Chytridiomycota	Chytridiales	*Polychytrium aggregatum* JEL109	in silico	neg			
Chytridiomycota	Chytridiales	*Polyphlyctis willoughbyi* PLAUS 26	in silico	neg			
Chytridiomycota	Spizellomycetes	*Powellomyces hirtus* Barr9B	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Powellomyces hirtus* BR81	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Powellomyces hirtus* JEL0540	in silico	neg			
Zoopagomycota	Kickxellomyocotina	*Ramicandelaber brevisporus* CBS 109374	in vivo	pos	3		
Zoopagomycota	Kickxellomyocotina	*Ramicandelaber longisporus* ARSEF 6175	in silico	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium globosum* JEL800	in vivo	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium* sp. JEL0864	in silico	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium* sp. JEL0881	in vivo	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium* sp. JEL0884	in vivo	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium* sp. JEL0917	in vivo	neg			
Mucoromycota	Glomeromycotina	*Rhizophagus diaphanum* MUCL 43196	in vivo	neg			
Mucoromycota	Glomeromycotina	*Rhizophagus intraradices*	in silico	pos		5	Branch 1
Mucoromycota	Glomeromycotina	*Rhizophagus irregularis*	in silico	pos		2	Branch 1
Mucoromycota	Glomeromycotina	*Rhizophagus* sp. strain HR1	in silico	pos		9	Branch 1 (4), Branch 2 (3), Branch 3

							(1), Unassigned (1)
Chytridiomycota	Rhizophlyctidales	*Rhizophlyctis rosea* JEL0318	in silico	pos		1	Branch 1
Chytridiomycota	Rhizophlyctidales	*Rhizophlyctis rosea* JEL0532	in vivo	neg			
Chytridiomycota	Rhizophlyctidales	*Rhizophlyctis rosea* JEL0564	in vivo	neg			
Chytridiomycota	Rhizophlyctidales	*Rhizophlyctis rosea* JEL0764	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Rhizophydium* sp. JEL0728	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Rhizophydium* sp. JEL0801	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Rhizophydium* sp. JEL0829	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Rhizophydium* sp. JEL0838	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Rhizophydium* sp. JEL0862	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Rhizophydium* sp. JEL0866	in vivo	neg			
Chytridiomycota	Rhizophydiales	*Rhizophydium* sp. MP83	in vivo	neg			
Mucoromycota	Mucoromycotina	*Rhizopus microsporus*	in vivo	neg			
Mucoromycota	Mucoromycotina	*Rhizopus oryzae*	in silico	neg			
Mucoromycota	Mucoromycotina	*Rhizopus stolonifer*	in silico	pos		2	Branch 1, Branch 4
Cryptomycota		*Rozella allomycis* CSF55	in silico	neg			
Chytridiomycota	Cladochytriales	*Septochytrium* sp. JEL0177	in silico	neg			
Zoopagomycota	Kickxellomyocotina	*Smittium commune* ARSEF 9245	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Spizellomyces* sp. JEL0132	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Spizellomyces* sp. JEL0210	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Spizellomyces* sp. JEL0361	in vivo	neg			
Mucoromycota	Mucoromycotina	*Syncephalastrum racemosum* NRRL 2495	in vivo	neg			
Zoopagomycota	Zoopagomycotina	*Syncephalis fuscata* S228	in vivo, in silico	neg, neg			
Zoopagomycota	Zoopagomycotina	*Syncephalis plumigaleata* NRRL S24	in silico	pos		3	Branch 1 (2), Branch 3 (1)
Chytridiomycota	Synchytriales	*Synchytrium endobioticum*	in silico	neg			

Chytridiomycota	Synchytriales	*Synchytrium microbalum* JEL517	in vivo	neg			
Mucoromycota	Mucoromycotina	*Thamnostylum piriforme* NRRL A-21589	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar arcticum* BR59	in silico	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar* sp. JEL0642	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar* sp. JEL0683	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar* sp. JEL0684	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar* sp. JEL0688	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar* sp. JEL0740	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar* sp. JEL0817	in vivo	neg			
Mucoromycota	Mucoromycotina	*Umbelopsis isabellina* AD 026	in vivo, in silico	neg, neg			
Mucoromycota	Mucoromycotina	*Umbelopsis nana* TLT 204	in vivo, in silico	pos, neg	2		
Mucoromycota	Mucoromycotina	*Umbelopsis ramanniana* AG	in silico	neg			
Mucoromycota	Mucoromycotina	*Umbelopsis vinacea* A-13231	in vivo	pos	2		
Chytridiomycota	Chytridiales	unknown JEL0085	in vivo	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium umbonatum* var. *sphaericum* JEL0516	in vivo	neg			
Chytridiomycota	Chytridiales	unknown JEL0546	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0547	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0614	in vivo	pos	1		
Chytridiomycota	Spizellomycetes	Unknown JEL0650	in vivo	neg			
Chytridiomycota	Cladochytriales	unknown JEL0793	in vivo	neg			
Chytridiomycota	Cladochytriales	*Chytriomyces hyalinus* JEL0795	in vivo	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium umbonatum* var. *sphaericum* JEL0796	in vivo	neg			
Chytridiomycota	Spizellomycetes	*Triparticalcar* sp. JEL0813	in vivo	neg			
Chytridiomycota	Chytridiales	*Rhizoclosmatium pessaminum* JEL0823	in vivo	neg			
Chytridiomycota	Chytridiales	unknown JEL0831	in vivo	neg			
Chytridiomycota	Chytridiales	unknown JEL0832	in vivo	neg			

Chytridiomycota	Chytridiales	unknown JEL0833	in vivo	neg			
Chytridiomycota	Chytridiales	unknown JEL0834	in vivo	pos	3		
Chytridiomycota	Chytridiales	unknown JEL0851	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0855	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0856	in vivo	neg			
Chytridiomycota	Chytridiales	unknown JEL0875	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0876	in vivo	neg			
Chytridiomycota	Chytridiales	unknown JEL0878	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0886	in vivo	neg			
Chytridiomycota	Monoblepharidales	unknown JEL0889	in vivo	neg			
Chytridiomycota	Monoblepharidales	unknown JEL0892	in vivo	neg			
Chytridiomycota	Cladochytriales	unknown JEL0904	in vivo	neg			
Chytridiomycota	Chytridiales	unknown JEL0906	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0911	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown JEL0915	in vivo	neg			
Chytridiomycota	Rhizophydiales	unknown MP49	in vivo	neg			
Chytridiomycota	Chytridiales	*Wheelerophlyctis interiexterior* JEL0857	in vivo	pos	1		
Chytridiomycota	Chytridiales	*Wheelerophlyctis interiexterior* JEL0885	in vivo	neg			
Zoopagomycota	Entomophthoromycotina	*Zoophthora radicans* ARSEF 6917	in vivo	pos			
Zoopagomycota	Entomophthoromycotina	*Zoophthora radicans* ATCC 208865/ARSEF 4784	in silico	pos		2	Branch 2, Branch 4
Chytridiomycota	Chytridiales	*Zopfochytrium polystomum* JEL0600	in vivo	pos	2		
Chytridiomycota	Chytridiales	*Zopfochytrium polystomum* WB228	in vivo, in silico	pos, pos	2	8	Branch 4

Supplemental Table 2.2 Collection information for isolates screened.

Isolate	Country	State/ Province	City	Location	Habitat	Year	Other Notes
Absidia californica NRRL 2967	USA	CA		Telegraph Peak, San Gabriel Mountain, California	Mouse dung		
Absidia corymbifera NRRL A-14836	Iceland				Volcanic soil		
Absidia ramosa NRRL 3180					peanut meal with aflatoxin		
Absidia repens NRRL 1336					soil		
Actinomucor elegans NRRL 3104	Taiwan		Taipei		Chinese cheese		
Allomyces sp. DJ02	USA	AL		Orange Beach	soil		
Allomyces sp. DJ05	USA	MS	Hattiesburg		soil		
Allomyces sp. DJ07	USA	MS	Meridian		soil		
Allomyces sp. WWM105							
Allomyces arbuscula Brit East Africa 2							
Allomyces arbuscula ATCC 10983							
Allomyces arbuscula Costa Rica 18	Costa Rica						
Allomyces arbuscula Cuba S 20	Cuba						
Allomyces arbuscula Mexico 57	Mexico						
Allomyces arbuscula Ohio 5	USA	OH					
Allomyces arbuscula Cuba S 22	Cuba						
Allomyces arbuscula FGSC48							
Allomyces javanicus Cuba S 12	Cuba						

Allomyces macrogynus ATCC_38327	Myanmar						
Allomyces moniliformis JMM01	India						
Allomyces javanicus India-4	India						
Allomyces macrogynus Australia-3	Australia						
Allomyces anomalus California-6	USA	CA					
Allomyces sp. BEA							
Allomyces arbuscula Burma 1F	Myanmar						
Allomyces arbuscula Belgian Congo 1	Democratic Republic of Congo						
Allomyces arbuscula Burma 1Db	Myanmar						
Allomyces arbuscula California-7	USA	CA					
Allomyces arbuscula CBS10.463							
Allomyces arbuscula Costa Rica 57A	Costa Rica						
Allomyces arbuscula Cuba S 28	Cuba						
Allomyces arbuscula Denmark 1	Denmark						
Allomyces arbuscula El Salvador 1	El Salvador						
Allomyces arbuscula Florida F2/F12	USA	FL					
Allomyces arbuscula North Carolina 2	USA	NC					
Allomyces arbuscula Philippines Isl 1	Philippines						
Allomyces javanicus California-1	USA	CA					

Allomyces sp. Australia CaC	Australia						
Allomyces sp. Fla 2	USA	FL					
Amylomyces rouxii NRRL 5866	Thailand						
Amylomyces rouxii NRRL A-11375							
Amylomyces rouxii NRRL A-25885					ragi		
Amylomyces rouxii NRRL A-26221	Phillipines		Bayombong				
Anaeromyces sp. NHY-2018							
Anaeromyces sp. S4							
Backusella circina 2446	USA	FL			soil with lichens		
Backusella ctenidia 6239	USA	CA		Death Valley	soil		
Basidiobolus sp. JELUM001							
Batrachochytrium dendrobatidis 423	Panama						
Batrachochytrium dendrobatidis Campana13	Brazil						
Batrachochytrium dendrobatidis Campana24	Brazil						
Batrachochytrium dendrobatidis CLFT044	Brazil						
Batrachochytrium dendrobatidis FMB03							
Batrachochytrium dendrobatidis NAF077	Brazil						
Batrachochytrium dendrobatidis NMBF05	Brazil						
Batrachochytrium dendrobatidis SAF014	Brazil						
Batrachochytrium dendrobatidis UM142	USA	MI					
Batrachochytrium salamandrivorans	The Netherlands		Bunderbos				

Benjaminiella poitrasii RSA 903							
Bifiguratus adelaidae AZ 501	USA	NC	Durham	Duke Forest	free air CO2 enrichment	2013	
Bifiguratus adelaidae TLT 265	USA	NC	Durham	Duke Forest	free air CO2 enrichment	2013	
Blakeslea trispora F986							
Blakeslea trispora NRRL 2456	Panama				soil		
Blakeslea trispora NRRL 2895	Panama			Albrook Field, Panama Canal Zone	soil		
Blastocladiella britanica JEL0711	USA	AK		Eagle Summit		2011	
Blastocladiella brittanica Barr214	England			Lake District National Park	1959		
Blyttiomyces sp. JEL0837							
Boealophlyctis nickersoniae WJD171	USA	AL				2011	Cotton field off I-65
Borealophlyctis nickersonii WJD170	USA	AL		Talladega National Forest	2012		
Catenaria sp. JEL0871						2015	
Catenaria sp. JEL748	USA	MI	South Haven	Spring Hill Pond	pond water	2012	
Catenaria sp. MP54							
Catenaria sp. MP55							
Catenaria sp. PL171							
Catenochytridium sp. JEL0775	USA	ME	Old Town	Pushaw Lake		2012	
Chaetocladium brefeldii NRRL 2508					horse dung		
Chaetocladium brefeldii NRRL 2343	USA	WI	near Saulk City	Ferry Bluff	Forest soil		
Chlamydoabsidia padenii NRRL 2977	USA	ID		University of Idaho	pea plant root rot		
Choanephora cucurbitarum NRRL 2744							

Chytridium lagenaria Arg66	Argentina	MZ			stream water		stream
Chytriomyces hyalinus JEL0117	USA	MI		Bryant's Bog		1992	
Chytriomyces hyalinus JEL0632	USA	ME	Orono	University Inn	soil	2009	
Chytriomyces sp. nov. MP71							
Circinella angarensis NRRL 2628	USA	CA	Claremont	San Gabriel Mountains	mouse dung		
Circinella minor NRRL A-13969	Australia	Adelaide			mouse dung		
Circinella muscae NRRL 1364					Gold mine soil at a depth of 600 meters at 40C		
Circinella naumovii NRRL 5846							
Cladochytrium sp. JEL0144	USA	ME		Mud Pond		1994	
Cladochytrium sp. JEL0153	USA	ME	Hampden	Woodland Pond	pond water	1994	
Cladochytrium sp. JEL0477	USA	ME	Augusta		pond water	1989	cattail pond in shallow granite quarry
Cladochytrium sp. JEL0479	USA	ME	Orono	University of Maine Campus	pond water	1990	cattail pond/garden pond
Cladochytrium sp. JEL0521	USA	ME		Mud Pond		2006	
Cladochytrium sp. JEL0592	USA	MI		Smith's Fen		2008	
Cladochytrium sp. JEL0704	USA	ME	Old Town	Mud Pond	pond water	2011	
Cladochytrium sp. JEL0714	USA	ME	Old Town	Pushaw Lake		2011	Cook's Point on Pushaw Lake
Cladochytrium sp. JEL0772	USA	ME	Orono		pond water	2012	pond or ditch on Godfrey Rd. by MBNA
Cladochytrium sp. JEL0861	USA	ME	Old Town	Mud Pond		2015	

Cladochytrium sp. JEL0893	USA	ME	Bucksport	Williams Pond		2016	
Cladochytrium sp. JEL0899	USA	ME	Bucksport	Williams Pond		2016	
Cladochytrium sp JEL0900	USA	ME	Old Town	Pushaw Lake		2016	Cook's Point on Pushaw Lake
Cladochytrium sp. JEL0903	USA	ME	Old Town	Pushaw Lake		2016	Cook's Point on Pushaw Lake
Cladochytrium tenue GHJ CCIBt 4013	Brazil	SAO	Cananéia	Ilha de Cardoso State Park	water and soil	2013	
Claroideoglomus etunicatum							
Coelomomyces lativattus CIRM-AVA-1							
Coemansia mojavensis RSA 71							
Coemansia reversa NRRL1564					seed, Brazil nut		
Coemansia spiralis RSA 1278							
Cokeromyces recurvatus NRRL 2243							
Conidiobolus antarcticus ARSEF 6913				Antarctica, Victoria Land	1999		
Conidiobolus coronatus ARSEF 9914	USA	VT				2010	
Conidiobolus coronatus NRRL 28638	Sweden						
Conidiobolus coronatus WDUM101	USA	MI	Ann Arbor	Nichol's Arboretum	leaf litter		
Conidiobolus coronatus WDUM102							
Conidiobolus thromboides FSU 785							
Cunninghamella sp. NRRL A-21271	Pakistan		Lyallpur		Insect, red bug in soil		
Dichotomocladium elegans RSA 919	USA	CA			rat dung		

Dissophora decumbens NRRL 22416	USA	RI			plant, ground-up Quercus (Oak) and Acer (Maple) leaves		
Dissophora ornata CBS 347.77	Colombia			Cordillera Central, Cauca en Huila, Parque Nacional del Puracé	soil, in mountain forest under Weinmannia, Clusia etc., alt. 3100 m		
Endochytrium sp. JEL0050	USA	ME	Orono			1989	
Endochytrium sp. JEL0386	USA	NY		Adirondack Mountains	2003		
Endochytrium sp. JEL0896	USA	ME	Bucksport	Williams Pond		2016	
Entomophaga maimaga ARSEF 7190	USA	PA					
Entomophthora muscae HHdFL130914-01							
Entophlyctis helioformis JEL0805	USA	ME	Old Town	Pushaw Lake		2013	
Entophlyctis sp. JEL0112	USA	ME	Corea	Corea Bog		1992	
Fimicolochytrium alabamae JEL0588	USA	ME	Orono	Longcore garden		2008	
Fimicolochytrium sp. JEL0733	USA	ME	Old Town	Witter Teaching and Research Center	2010		
Furculomyces boomerangus NRRL 9021	Australia	Tasmania				1987	
Gaertneriomyces sp. JEL0119	USA	ME			manure	1993	horse manure, water, pollen
Gaertneriomyces sp. JEL0628	USA	ME	Orono	University Inn	soil	2009	
Gaertneriomyces sp. JEL0657	USA	ME	Old Town	Witter Teaching and Research Center	manure	2010	manure from bunker at Witter Farm
Gaertneriomyces sp. JEL0662	USA	ME	Old Town	Witter Teaching and Research Center	manure	2010	manure from bunker at Witter Farm
Gaertneriomyces sp. JEL0698	USA	AK	Fairbanks	University of Alaska Fairbanks	manure	2011	musk ox dung

				Large Animal Farm			
Gaertneriomyces sp. JEL0699	USA	AK	Fairbanks	University of Alaska Fairbanks Large Animal Farm	manure	2011	caribou dung
Gaertneriomyces sp. JEL0738	USA	ME	Old Town	Witter Teaching and Research Center	2010		
Gaertneriomyces semiglobifer Barr 43	Germany			Baltic Sea	beach sand		
Geosiphon pyriformis							
Geranomyces michiganensis JEL0563	USA	MI	Alpena		soil	2007	Al's camp dirt
Geranomyces variabilis Barr350	Canada	NB		Kouchibougnacis National Park	soil	1978	
Geranomyces variabilis JEL 559	USA	MI	Alpena		manure	2007	old horse manure
Geranomyces variabilis JEL0542	USA	ME		Salmon Pond		2007	
Geranomyces variabilis JEL0557	USA	MI	Alpena		manure	2007	Al's camp, aged cowpie
Geranomyces variabilis JEL0566	USA	MI	Alpena		manure	2007	old horse manure
Geranomyces variabilis JEL0567	USA	MI	Alpena		manure	2007	old horse manure
Geranomyces variabilis KP27	Scotland		Rothes			2006	
Geranomyces variabilis KP31	USA	VA		Goshen Wildlife Management Area	soil	2007	sandy soil
Geranomyces variabilis MP0003	USA	NC	Chapel Hill		soil	1968	lawn soil
Geranomyces variabilis MP0004	USA	AL	Northport	Lake Lurleen State Park	soil	2006	forest soil
Gigaspora margarita BEG34							
Gigaspora rosea							
Gilbertella persicaria NRRL A-13613	South Africa				soil		

Gilbertella persicaria var. *persicaria* CBS 190.32	USA	NY			Prunus persica fruit		
Globomyces pollinis-pini Arg68	Argentina	BA	Buenos Aires				
Gonapodya sp. JEL0183	USA	ME	Orono		duckweed	1996	MBNA Orono Duck Pond
Gongronella butleri NRRL 1340							
Gongronella butleri NRRL A-23795					ant fungus garden		
Gongronella butleri C1D							
Gorgonomyces sp. Arg29	Argentina	CN			stream water		stream
Gorgonomyces haynaldii MP0057							
Halteromyces radiatus CBS 162.75 T	Australia	Queensland			mud from mangrove forest, contaminated with effluent	1972	
Harpochytrium sp. JEL0705	USA	ME	Old Town	Mud Pond		2011	
Helicostylum pulchrum RSA 2064							
Hesseltinella vesiculosa NRRL 3301	Brazil				soil, a paddy field		
Hesseltinella vesiculosa NRRL 34							
Hyaloraphidium curvatum JEL0383	South Africa				freshwater		
Kappamyces sp. JEL0680	USA	ME	Old Town	Witter Teaching and Research Center	soil	2010	lawn soil
Kickxella alabastrina RSA 675	USA	CA	Claremont	Evey Canyon, 45 miles north-east of Claremont, California	rat dung		
Kirkomyces cordense RSA 1222							
Legeriosimilis sp. ARSEF 9066	Canada	Ontario				2004	

Linderina pennispora ATCC12442/NRRL2237	Liberia		Monrovia				
Lobosporangium transversale NRRL 3116	USA	NV	Virginia City		soil beneath Purshia tridentata		
Lobulomyces angularis JEL0522	USA	TN	Athens	Mayfield Dairy Farm			
Lobulomyces poculatus JEL0511	USA	ME	Old Town	Mud Pond		2005	
Martensiomyces pterosporus CBS 209.56	Zaire				forest soil	1955	
Mortierella alpina							
Mortierella ambigua NRRL 28271							
Mortierella capitata AV 005	USA	Puerto Rico				2016	
Mortierella echinospaera NRRL 1233	Netherlands	North Holland	Aalsmeer				
Mortierella elongata NRRL 5513	USA	GA			golf course soil		
Mortierella elongata NVP 64	USA	MI	Jackson			2015	
Mortierella epicladia AD 058	USA	MI	Laingsburg	Sleepy Hollow State Park	Rhizosphere/Roots of Spruce and Pine	2015	
Mortierella gamsii AM 1032							
Mortierella humilis PMI 1414	USA	MA				2009	
Mortierella hyalina NRRL 6427	USA	NY			Fungus, Hypoxylon deustum		
Mortierella minutissima AD051							
Mortierella multidivaricata NRRL 6456							
Mortierella multidivaricata RSA 2152							
Mortierella nov. sp. GBAus 27b	USA	IL				2013	

Mortierella polycephala KOD 948	USA	IL			soil	2012	
Mortierella selenospora CBS 811.68	The Netherlands				mushroom compost	1968	
Mortierella selenospora KOD 1015	USA	IL			cave wall	2013	
Mortierella sp. GBAus 30	Australia				soil		
Mortierella sp. JEL0843	USA	ME	Old Town	Mud Pond	soil	2015	
Mortierella sp. JEL0858	USA	ME	Old Town	Mud Pond		2015	
Mortierella sp. JEL0860	USA	ME	Old Town	Mud Pond		2015	
Mortierella verticillata NRRL 6337	United Kingdom						
Mortierella wolfii NRRL 28640							
Mucor circinelloides							
Mucor griseo-ochraceus var. *minuta* NRRL 3246							
Mucor heimalis WDUM104 777					soil pH 7		
Mucor hiemalis f. hiemalis NRRL 3624	USA	MI					
Mucor irregularis C3B							
Mucor lahorensis NRRL 6592	Pakistan						
Mucor mousanensis NRRL 3105							
Mucor pakistanicus NRRL 6589	Pakistan						
Mucor petrinsularius var. *echinosporus* NRRL 3141							
Mucor petrinsularius var. *ovalisporus* NRRL 2536				Clinical isolate from a human ear			
Mucor ramosissimus NRRL 3042	Uruguay				clinical isolate		
Mucor recurvus NRRL 2358							
Mucor rouxii NRRL 3367							

Mucor rouxii NRRL A-11340	Indonesia	Java			food, ragi		
Mucor rouxii NRRL A-11341	Indonesia	Java			food, ragi		
Mucorales sp. UM774	USA	MI	Ann Arbor	Nichol's Arboretum	leaf litter	2016	
Mycocladus corymbifer NRRL 1309	Ghana				Theobroma cacao		
Mycotypha africana NRRL 2978	Zimbabwe			South of Umtali	soil		
Mycotypha microspora NRRL 684							
Neocallimastix californiae	USA	IA			wood		
Neocallimastix frontalis 27							
Nephrochytrium aurantium JEL0909	USA	WA	Kirkland	Edith Moulton Park	2017		
Nowakowskiella sp. JEL0774	USA	ME	Old Town	Pushaw Lake		2012	
Nowakowskiella sp. JEL0785	USA	ME				2015	red maple leaves with artificial stream water
Obelidium sp. JEL0802	USA	ME	Old Town	Pushaw Lake		2013	
Operculomyces laminatus JEL223	USA	ME	Orono	Longcore garden	soil	1998	garden soil
Orpinomyces sp. C1A							
Orpinomyces joyonii SG4							
Paraglomus brasilianum							
Paramicrosporidium saccamoebae	Germany			Luneburg Heath	pond water		
Paranamyces uniporus JEL0695	USA	ME	Orono	Longcore garden	soil	2011	
Paranomyces uniporus WJD150	USA	AL	Tuscaloosa	Riverside Pond	soil	2011	
Paranomyces uniporus WJD158	USA	AL	Oxford		soil	2011	

Paranomyces uniporus WJD193	USA	OH	Bath Township	Bath Nature Preserve	2011	tamarack bog	
Paraphysoderma sedebokerense JEL0821	USA					2014	
Paraphysoderma sedebokerense JEL0847							
Parasitella parasitica NRRL 2501							
Phascolomyces articulosus RSA 2281							
Phycomyces blakesacanus UM 175							
Phycomyces blakesleeanus							
Phycomyces nitens S609							
Phylctochytrium bullatum JEL0754	USA	MI	South Haven	Spring Hill Pond		2012	
Phylctochytrium planicorne JEL0894	USA	ME	Bucksport	Williams Pond		2016	shore
Physocladia obscura JEL0137	USA	MI		Bryant's Bog		1992	
Piptocephalis cylindrospora RSA 2659							
Piromyces rhizinflatus YM600							
Piromyces finn							
Polychytrium aggregatum JEL109	USA	MI		Bryant's Bog		1992	
Polyphlyctis willoughbyi PLAUS 26	Australia	NSW	Morton National Park Clyde River				
Powellomyces hirtus Barr9B	Canada	ON	Ottawa		soil	1966	garden soil
Powellomyces hirtus BR81							
Powellomyces hirtus JEL0540	USA	CA	Pacific Grove		stream water	2007	

Ramicandelaber brevisporus CBS 109374	Japan		Tokyo	Hachijyo Island	soil	1999	
Ramicandelaber longisporus ARSEF 6175	China		Beijing			1992	
Rhizoclosmatium globosum JEL800	USA	ME	Old Town	Mud Pond		2013	
Rhizoclosmatium sp. JEL0864	USA	ME	Old Town	Mud Pond		2015	rozella host
Rhizoclosmatium sp. JEL0881	USA	ME	Orono	Littlefield Ornamental Garden	pond water	2016	small pond
Rhizoclosmatium sp. JEL0884	USA	ME	Old Town	Mud Pond		2016	
Rhizoclosmatium sp. JEL0917	USA	ME	Bucksport	Williams Pond		2017	
Rhizophagus diaphanum MUCL 43196							
Rhizophagus intraradices							
Rhizophagus irregularis							
Rhizophagus sp. strain HR1							
Rhizophlyctis rosea JEL0318	USA	GA	Athens	Georgia Botanical Gardens	soil		
Rhizophlyctis rosea JEL0532	USA	MI	Alpena		soil	2006	Al's camp, dry woods soil
Rhizophlyctis rosea JEL0564	USA	LA	Baton Rouge	Lousiana State University	tree detritus	2007	"tree detritus, moss on oaks"
Rhizophlyctis rosea JEL0764	USA	ME	Milford		aquatic grass	2012	pond
Rhizophydium sp. JEL0728	USA	ME		Sewall Tract Forest Path	2011		
Rhizophydium sp. JEL0801	USA	ME	Old Town	Pushaw Lake		2013	
Rhizophydium sp. JEL0829							
Rhizophydium sp. JEL0838							
Rhizophydium sp. JEL0862	USA	ME	Old Town	Mud Pond		2015	

Rhizophydium sp. JEL0866	USA	ME	Old Town	Mud Pond		2015	
Rhizophydium sp. MP83	USA	AL	Tuscaloosa	Black Warrior River	water and soil		
Rhizopus microsporus							
Rhizopus oryzae							
Rhizopus stolonifer							
Rozella allomycis CSF55	USA	MS	Hattiesburg		soil		
Septochytrium sp. JEL0177	USA	ME	Wales			1995	"dry detritus from edge of lake in Wales"
Smittium commune ARSEF 9245	USA	KS				1993	
Spizellomyces sp. JEL0132	USA	ME	Orono	University of Maine Campus	soil	1993	
Spizellomyces sp. JEL0210	USA	PR			soil	1998	
Spizellomyces sp. JEL0361	Australia	QL					
Syncephalastrum racemosum NRRL 2495							
Syncephalis fuscata S228							
Syncephalis plumigaleata NRRL S24	USA	MA		Smith College under large Gingko	soil		
Synchytrium endobioticum							
Synchytrium microbalum JEL517	USA	ME	Old Town	Mud Pond		2006	
Thamnostylum piriforme NRRL A-21589							
Triparticalcar arcticum BR59	Pakistan		Lyallpur		rabbit dung		
Triparticalcar sp. JEL0642					manure	2010	horse manure
Triparticalcar sp. JEL0683	USA	ME	Old Town	Witter Teaching and Research Center	manure	2010	

Triparticalcar sp. JEL0684	USA	ME	Old Town	Witter Teaching and Research Center	manure	2010	
Triparticalcar sp. JEL0688	USA	ME	Old Town	Witter Teaching and Research Center	manure	2010	
Triparticalcar sp. JEL0740	USA	ME	Old Town	Witter Teaching and Research Center	2010		
Triparticalcar sp. JEL0817	USA	VA		Assateague Island	2014		
Umbelopsis isabellina AD 026	USA	AZ					Collected near Humphrey's Peak
Umbelopsis nana TLT 204							
Umbelopsis ramanniana AG							
Umbelopsis vinacea A-13231							
unknown JEL0085	USA	ME	Orono	University of Maine Campus	1991		
Rhizoclosmatium umbonatum var. *sphaericum* JEL0516	USA	ME	Old Town	Mud Pond		2006	
unknown JEL0546	USA	CA	Pacific Grove	Asilomar Stream	2007		
unknown JEL0547	USA	CA	Pacific Grove	Asilomar Stream	2007		
unknown JEL0614	USA	UT	Snowbird	Hidden Peak	soil	2009	"mud under snow"
Unknown JEL0650	USA	ME				2010	
unknown JEL0793	USA	ME	Orono	Pond by MBNA		2013	"with grass leaves Godfrey Rd. ditch"
Chytriomyces hyalinus JEL0795	USA	ME		Little Deer Isle	soil	2013	
Rhizoclosmatium umbonatum var. *sphaericum* JEL0796	USA	ME	Old Town	Mud Pond		2013	

Triparticalcar sp. JEL0813	USA	VA		Assateague Island	2014		
Rhizoclosmatium pessaminum JEL0823	USA	ME	Old Town	Mud Pond		2014	
unknown JEL0831							
unknown JEL0832							
unknown JEL0833							
unknown JEL0834							
unknown JEL0851	USA	ME	Old Town	Mud Pond		2015	
unknown JEL0855	USA	ME		Aroostook River		2015	
unknown JEL0856	USA	ME	Old Town	Mud Pond		2015	
unknown JEL0875	USA	ME	Orono	Stillwater River	soil	2015	
unknown JEL0876	USA	PA		Black Moshannon Bog	2015		
unknown JEL0878	USA	ME	Old Town		frog	2016	Dave Thompson's Pond, dead frog toe webbing
unknown JEL0886	USA	ME	Orono	Littlefied Ornamental Garden			
unknown JEL0889	USA	MA	West Tisbury	North Dewarts Pond	2016		
unknown JEL0892	USA	ME	Bucksport	Williams Pond		2016	shore
unknown JEL0904	USA	ME	Old Town	Pushaw Lake		2016	
unknown JEL0906	USA	WA	Kirkland	Edith Moulton Park	soil	2017	"mud"
unknown JEL0911	USA	ME	Orono	Pond by MBNA		2017	"Godfrey Road Pond"
unknown JEL0915						2017	
unknown MP49	USA	AL	Tuscaloosa	Black Warrior River			
Wheelerophlyctis interiexterior JEL0857	USA	ME	Old Town	Mud Pond		2015	

Wheelerophlyctis interiexterior JEL0885	USA	ME	Old Town			2016	
Zoophthora radicans ARSEF 6917	Argentina	Buenos Aires	Buenos Aires		on Lycoperisoc esculentum, tomato in covered greenhouses.	2001	
Zoophthora radicans ATCC 208865/ARSEF 4784							
Zopfochytrium polystomum JEL0600	USA	ME		Bear Brook Watershed	2008	"water and submerged leaves"	
Zopfochytrium polystomum WB228	USA	AL		Lake Lurleen		2005	"submerged shoreline vegetation, mud, and water"

Supplemental Table 2.3 Transcriptomes retrieved from SRA database.

Isolate	**SRR**	**Citation**
Absidia repens NRRL 1336	SRR7974509	Mondo SJ et al., 2017 doi: 10.1038/ng.3859
Allomyces macrogynus ATCC_38327	SRR343045	unpublished
Amylomyces rouxii NRRL 5866	SRR8529687	unpublished
Anaeromyces sp. NHY-2018	SRR7819341	unpublished
Anaeromyces sp. S4	SRR4063399	Haitjema CH et al., 2017 doi: 10.1038/nmicrobiol.2017.87
Batrachochytrium dendrobatidis 423	SRR2719455	Ellison, AR et al., 2017 doi: 10.1534/g3.116.035873
Batrachochytrium dendrobatidis CLFT044	SRR5988742	McDonald, CA et al., 2019 doi: 10.1016/j.funbio.2019.10.008
Batrachochytrium salamandrivorans	SRR3706726	Farrer, RA et al. 2017 doi: 10.1038/ncomms14742
Benjaminiella poitrasii RSA 903	SRR6942796	unpublished
Blakeslea trispora F986	SRR8238938	unpublished
Blakeslea trispora NRRL 2456	SRR6049667	unpublished
Blastocladiella britanica JEL0711	SRR6057017	unpublished
Catenaria sp. PL171	SRR424218	Mondo SJ et al., 2017 doi: 10.1038/ng.3859
Chaetocladium brefeldii NRRL 2343	SRR7975436	unpublished
Chlamydoabsidia padenii NRRL 2977	SRR6057190	unpublished
Choanephora cucurbitarum NRRL 2744	SRR8238940	unpublished
Chytridium lagenaria Arg66	SRR6057011	unpublished
Chytriomyces hyalinus JEL0632	SRR8189667	unpublished
Chytriomyces sp. nov. MP71	SRR6056998	unpublished
Claroideoglomus etunicatum	DRR041158	unpublished
Coelomomyces lativattus CIRM-AVA-1	SRR5504042	Ahrendt, SA. 2015 ProQuest ID: Ahrendt_ucr_0032D_12303
Coemansia mojavensis RSA 71	SRR7140835	unpublished
Coemansia spiralis RSA 1278	SRR6056689	unpublished
Cokeromyces recurvatus NRRL 2243	SRR6056974	unpublished
Conidiobolus thromboides FSU 785	SRR4052359	unpublished
Dichotomocladium elegans RSA 919	SRR6256436	unpublished
Dissophora ornata CBS 347.77	SRR8238935	unpublished
Entomophaga maimaga ARSEF 7190	SRR9001799	unpublished
Entomophthora muscae HHdFL130914-01	SRR5506701	Nibert, ML et al. 2019 doi: 10.3390/v11040351
Entophlyctis helioformis JEL0805	SRR6057018	unpublished
Gaertneriomyces semiglobifer Barr 43	SRR6056997	unpublished
Geosiphon pyriformis	SRR6363035	unpublished
Geranomyces variabilis JEL 559	SRR8534491	unpublished
Gigaspora margarita BEG34	SRR1659851	Salvioli, A 2010 doi: 10.1111/j.1462-2920.2010.02246.x
Gigaspora rosea	SRR1979254	Tang, N 2016 doi: 10.3389/fmicb.2016.00233

Gilbertella persicaria var. *persicaria* CBS 190.32	SRR6056754	unpublished
Globomyces pollinis-pini Arg68	SRR6057014	unpublished
Gongronella butleri C1D	SRR8303828	unpublished
Gorgonomyces haynaldii MP0057	SRR8267450	unpublished
Halteromyces radiatus CBS 162.75 T	SRR7476973	unpublished
Helicostylum pulchrum RSA 2064	SRR8240038	unpublished
Hesseltinella vesiculosa NRRL 3301	SRR4063264	Mondo SJ et al., 2017 doi: 10.1038/ng.3859
Hyaloraphidium curvatum JEL0383	SRR7517569	unpublished
Kickxella alabastrina RSA 675	SRR6057250	unpublished
Kirkomyces cordense RSA 1222	SRR6943043	unpublished
Linderina pennispora ATCC 12442/NRRL 2237	SRR3439779	Mondo SJ et al., 2017 doi: 10.1038/ng.3859
Lobosporangium transversale NRRL 3116	SRR8840873	Mondo SJ et al., 2017 doi: 10.1038/ng.3859
Martensiomyces pterosporus CBS 209.56	SRR4125806	unpublished
Mortierella alpina	SRR1638091	Wang, L et al., 2011 doi: 10.1371/journal.pone.0028319
Mortierella elongata (wildtype)	SRR6225696	Uehling, J et al., 2017 doi: 10.1111/1462-2920.13669
Mortierella elongata	SRR6225691	Uehling, J et al., 2017 doi: 10.1111/1462-2920.13669
Mortierella gamsii AM 1032	SRR5506997	unpublished
Mortierella humilis PMI 1414	SRR6256461	unpublished
Mortierella minutissima AD051	SRR6256835	unpublished
Mortierella multidivaricata RSA 2152	SRR6256450	unpublished
Mortierella nov. sp. GBAus 27b	SRR6256463	unpublished
Mortierella selenospora CBS 811.68	SRR6256464	unpublished
Mortierella verticillata NRRL 6337	SRR343048	unpublished
Mucor circinelloides	SRR9009727	Navarro-Mendoza MI et al., 2019 doi: 10.1016/j.cub.2019.09.024
Mucor irregularis C3B	SRR8992276	Barata, RR et al., 2019 doi: 10.1128/MRA.00503-19
Mycotypha africana NRRL 2978	SRR6049697	unpublished
Neocallimastix californiae	SRR5296032	Solomon, KV et al., 2016 doi: 10.1126/science.aad1431
Neocallimastix frontalis 27	SRR6829473	Gruninger, RJ et al., 2018 doi: 10.3389/fmicb.2018.01581
Obelidium sp. JEL0802	SRR5506997	unpublished
Operculomyces laminatus JEL223	SRR5579326	unpublished
Orpinomyces C1A	SRR2033890	Couger, MB et al., 2015 doi: 10.1186/s13068-015-0390-0
Orpinomyces joyonii SG4	SRR6829438	Gruninger, RJ et al., 2018 doi: 10.3389/fmicb.2018.01581
Paraglomus brasilianum	SRR5279411	Beaudet, D et al., 2018 doi: 10.1093/dnares/dsx051
Paraphysoderma sedebokerense JEL0821	SRR6942797	unpublished
Parasitella parasitica NRRL 2501	SRR8840827	unpublished
Phascolomyces articulosus RSA 2281	SRR6056973	unpublished
Phycomyces blakesleeanus	SRR9002717	unpublished
Phycomyces nitens S609	SRR6256453	unpublished

Piptocephalis cylindrospora RSA 2659	SRR5506707	Ahrendt, SR et al., 2018 doi:10.1038/s41564-018-0261-0
Piromyces rhizinflatus YM600	SRR6829441	Gruninger, RJ et al., 2018 doi: 10.3389/fmicb.2018.01581
Piromyces finn	SRR5487626	Haitjema CH et al., 2017 doi: 10.1038/nmicrobiol.2017.87
Polychytrium aggregatum JEL109	SRR9001797	unpublished
Powellomyces hirtus BR81	SRR7517594	unpublished
Ramicandelaber brevisporus CBS 109374	SRR4125814	unpublished
Rhizoclosmatium globosum JEL800	SRR5506698	Mondo SJ et al., 2017 doi: 10.1038/ng.3859
Rhizophagus diaphanum MUCL 43196	SRR916888	Tisserant E et al., 2013 doi: 10.1073/pnas.1313452110
Rhizophagus intraradices	SRR915897	Tisserant E et al., 2013 doi: 10.1073/pnas.1313452110
Rhizophagus irregularis	SRR7828199	Nuepane, A et al., 2018 doi:10.3390/v10120707
Rhizophagus sp. strain HR1	DRR017631	Kikuchi, Y et al., 2014 doi: 10.1111/nph.12937
Rhizopus microsporus	SRR4489400	Lastovetsky OA, et al., 2016 doi:10.1073/pnas.1615148113
Rhizopus oryzae	SRR2104505	unpublished
Rhizopus stolonifer	SRR8191390	Petrasch A, et al., 2019 doi: 10.3389/fpls.2019.00223
Syncephalastrum racemosum NRRL 2495	SRR8843209	Mondo SJ et al., 2017 doi: 10.1038/ng.3859
Syncephalis fuscata S228	SRR6053271	unpublished
Syncephalis plumigaleata NRRL S24	SRR6050220	unpublished
Triparticalcar arcticum BR59	SRR7517624	unpublished
Umbelopsis isabellina AD 026	SRR6942795	unpublished
Umbelopsis nana TLT 204	SRR6256849	unpublished
Umbelopsis ramanniana AG	SRR5487448	unpublished
Zoophthora radicans ATCC 208865/ARSEF 4784	SRR8189665	unpublished
Zopfochytrium polystomum WB228	SRR7141106	unpublished

Supplemental Table 2.4 GenBank accessions of reference viral sequences.

Virus label	GenBank accession
Acremonium sclerotigenum ourmia-like virus 1	QDB75006.1
Actinidia virus X	YP_009186834.1
Aedes angustivittatus narnavirus	QBA55490.1
Agaricus bisporus mitovirus 1	AQM32767.1
Allium cepa amalgavirus 1	YP_009447919.1
Alternaria alternata botybirnavirus 1	BBH54877.1
Alternaria arborescens mitovirus 1	YP_009270635.1
Alternaria arborescens victorivirus 1	YP_009553478.1
Anthoxanthum odoratum amalgavirus 2	DAB41668.1
Anthurium mosaic-associated virus	YP_009667023.1
Artichoke mottled crinkle virus	NP_039808.1
Asparagus virus 3	AIL23155.1
Aspergillus flavus partitivirus 1	QDE53634.1
Aspergillus homomorphus totivirus 1	AZT88629.1
Aspergillus ochraceous virus	ABC86749.1
Atkinsonella hypoxylon virus	NP_604475.1
Barns Ness breadcrumb sponge tombus-like virus 3	ASM94017.1
Barns Ness serrated wrack narna-like virus 4	ASM94100.1
Bean yellow mosaic virus	CDO67696.1
Beauveria bassiana RNA virus 1	AKC57301.1
Beihai barnacle virus 10	YP_009333179.1
Beihai barnacle virus 14	APG78182.1
Beihai narna-like virus 13	YP_009333241.1
Beihai narna-like virus 14	YP_009333153.1
Beihai narna-like virus 24	YP_009333245.1
Beihai narna-like virus 25	APG76998.1
Beihai tombus-like virus 6	APG76145.1
Beihai tombus-like virus 7	YP_009337688.1
Bipolaris maydis chrysovirus 1	ARM36035.1
Blueberry latent virus	BBH51573.1
Botryosphaeria dothidea chrysovirus 1	AGZ84312.1
Botryosphaeria dothidea victorivirus 2	QBA82443.1
Botrytis porri botybirnavirus 1	YP_006390636.1
Botrytis virus F	CFS87145.1
Botrytis virus X	NP_932306.1
Brassica campestris chrysovirus 1	YP_009667006.1
Brunton virus	QED21526.1
Bundaberg bee virus 5	AWK77861.1
Cactus virus X	BAU68240.1

Camponotus yamaokai virus	YP_009143313.1
Carnation cryptic virus 3	BBM06288.1
Carnation Italian ringspot virus	0ALJ30184.1
Cassava alphaflexivirus	AHA91819.1
Cassava virus C	ACI03053.1
Ceratobasidium mycovirus-like	AOX47586.1
Changjiang narna-like virus 5	APG77107.1
Changjiang tombus-like virus 1	YP_009337122.1
Cladosporium cladosporioides ourmia-like virus 1	QDB74999.1
Cladosporium cladosporioides ourmia-like virus 2	QDB75008.1
Colletotrichum caudatum totivirus 1	AZT88631.1
Colletotrichum fructicola chrysovirus 1	YP_009551629.1
Colletotrichum gloeosporioides ourmia-like virus 1	QDW80875.1
Colletotrichum zoysiae totivirus 1	AZT88637.1
Combu double-strand RNA mycovirus	QAB47444.1
Combu positive-strand RNA mycovirus	QAB47442.1
Coquillettidia venezuelensis narnavirus 1	QBA55488.1
Cryphonectria parasitica bipartite mycovirus 1	YP_007985675.1
Cryphonectria parasitica mitovirus 1-NB631	NP_660174.1
Cucumber necrosis virus	NP_040953.2
Culex Hubei-like virus	AXQ04814.1
Curvularia thermal tolerance virus	ALO61398.1
Curvularia virus 2	ALO61390.1
Darwin bee virus 4	AWK77851.1
Deformed wing virus	AKE50879.1
Delisea pulchra totivirus IndA	AMB17478.1
Diatom colony associated dsRNA virus 10	YP_009552793.1
Diatom colony associated dsRNA virus 11	YP_009552795.1
Diatom colony associated virus-Like RNA Segment 6	BAU79526.1
Donkey orchid virus A	YP_007969892.1
dsRNA virus environmental sample	AJT39579.1
Dumyat virus	QAY29251.1
Eggplant mosaic virus	NP_040968.1
Eggplant mottled crinkle virus	YP_008999611.1
Eimeria tenella RNA virus 1	YP_009115500.1
Entomophthora muscae mitovirus 1	QCF24445.1
Entomophthora muscae mitovirus 2	QCF24461.1
Entomophthora muscae mitovirus 3	DAC76942.1
Entomophthora muscae mitovirus 7	DAC76946.1
Erigeron breviscapus amalgavirus 1	YP_009552087.1
Erysiphe necator mitovirus 2	YP_009465716.1

Escherichia virus FI	YP_009208148.1
Festuca pratensis amalgavirus 2	YP_009553344.1
Fungal totivirus MpPl	ALD89107.1
Fusarium graminearum dsRNA mycovirus-4	YP_003288790.1
Fusarium graminearum mycotymovirus 1	YP_009553357.1
Fusarium poae mitovirus 4	YP_009272901.1
Fusarium poae narnavirus 1	YP_009272902.1
Fusarium poae victorivirus 1	YP_009272905.1
Giardia canis virus	ABB36743.1
Grapevine associated narnavirus-1	YP_009182162.1
Grapevine Syrah virus 1	AKZ17760.1
Gremmeniella abietina RNA virus 6	AIU98624.1
Gremmeniella abietina RNA virus L1	AAK11656.1
Gremmeniella mitovirus	CCD32685.2
Helicobasidium mompa mitovirus 1-18	BAD72871.1
Heterobasidion mitovirus 2	QED55404.1
Heterobasidion RNA virus 6	AHA82551.1
Hortaea werneckii totivirus 1	AZT88647.1
Hosta virus X	AFM77886.1
Hubei arthropod virus 1	YP_009336629.1
Hubei chryso-like virus 1	ASA47520.1
Hubei mosquito virus 1	YP_009337088.1
Hubei narna-like virus 13	YP_009337805.1
Hubei partiti-like virus 41	APG78238.1
Hubei partiti-like virus 59	APG78262.1
Hubei picorna-like virus 39	YP_009336612.1
Hubei picorna-like virus 40	YP_009336539.1
Hubei tombus-like virus 12	YP_009336735.1
Hubei toti-like virus 18	YP_009336932.1
Infectious flacherie virus	ABO09751.1
Lactarius rufus RNA virus 1	AMK47912.2
Lactarius tabidus RNA virus 1	AMK47915.3
Lampyris noctiluca toti-like virus 1	QBP37034.1
Leptomonas Narna-like virus 1	ATI23585.1
Leptopilina boulardi Toti-like virus	YP_009072448.1
Leshenault partiti-like virus	YP_009388588.1
Lettuce necrotic stunt virus	AFM91097.1
Lily virus X	QBQ82430.1
Linepithema humile toti-like virus 1	AXA52555.1
Luckshill virus	AWA82251.1
Macrophomina phaseolina mitovirus 1	ALD89100.1

Magnaporthe oryzae ourmia-like virus	YP_009667033.1
Magnaporthe oryzae partitivirus 1	APP18151.1
Magnaporthe oryzae virus 2	BBG92298.1
Maize chlorotic mottle virus	QDZ17053.1
Maize necrotic streak virus	YP_459920.2
Mitovirus AEF-2013	AGW51760.1
Mitovirus sp.	QDH90007.1
Mitovirus sp.	QDH89786.1
Mitovirus sp.	QDH87697.1
Mitovirus sp.	QDH89590.1
Mitovirus sp.	QDH88495.1
Mitovirus sp.	QDH87062.1
Mitovirus sp.	QDH89392.1
Mitovirus sp.	QDH90841.1
Mitovirus sp.	QDH89715.1
Mitovirus sp.	QDH87228.1
Mitovirus sp.	QDH88811.1
Mitovirus sp.	QDH87897.1
Mitovirus sp.	QDH86892.1
Mitovirus sp.	QDH90147.1
Mitovirus sp.	QDH91356.1
Mitovirus sp.	QDH88943.1
Mitovirus sp.	QDH88789.1
Mitovirus sp.	QDH90063.1
Mitovirus sp.	QDH89821.1
Mitovirus sp.	QDH88391.1
Mitovirus sp.	QDH91432.1
Mitovirus sp.	QDH87779.1
Mitovirus sp.	QDH88475.1
Mitovirus sp.	QDH88042.1
Mitovirus sp.	QDH89466.1
Mitovirus sp.	QDH91557.1
Mitovirus sp.	QDH88411.1
Mitovirus sp.	QDH91055.1
Mitovirus sp.	QDH86747.1
Mitovirus sp.	QDH87116.1
Mitovirus sp.	QDH91042.1
Mitovirus sp.	QDH90477.1
Mitovirus sp.	QDH88853.1
Mitovirus sp.	QDH86661.1
Mitovirus sp.	QDH89765.1

Mitovirus sp.	QDH89904.1
Moriarty virus	QED21524.1
Moroccan pepper virus	YP_009037606.1
Mycoreovirus 3	YP_392478.1
Myriodontium keratinophilum bipartite virus 1	AYP71809.1
Mytcor virus 1	AYN75557.1
narna-like virus 6	YP_009333214.1
Neofusicoccum parvum mitovirus 1	QDB74992.1
Nerine virus X	YP_446992.1
Nigrospora oryzae victorivirus 2	AZP53926.1
Oat blue dwarf virus	NP_044447.1
Ochlerotatus-associated narna-like virus 1	AGW51766.2
Ochlerotatus-associated narna-like virus 2	AGW51768.2
Odontoglossum ringspot virus	AAS87224.1
Ophiocordyceps sinensis mitovirus 1	AZT88623.1
Ophiostoma mitovirus 3a	NP_660176.1
Ourmia melon virus	ACF16360.1
Panax notoginseng virus A	YP_009225665.1
Partitiviridae sp.	QDH89651.1
Pea early-browning virus	AAA46820.1
Peach virus T	ARI47200.1
Penicillium aurantiogriseum bipartite virus 1	YP_009182335.1
Penicillium citrinum ourmia-like virus 1	AYP71797.1
Penicillium miczynskii RNA virus 1	QDB74980.1
Penicillium stoloniferum virus S	AAN86834.2
Pepper ringspot virus	NP_620033.1
Pericornia byssoides totivirus 1	QDB74982.1
Persea americana chrysovirus	YP_009666328.1
Phalaenopsis equestris amalgavirus 1	YP_009552083.1
Phomopsis longicolla totivirus 1	ALD89108.1
Phomopsis vexans RNA virus	YP_009115492.1
Phytophthora infestans RNA virus 4	YP_009241365.1
Pinus patula amalgavirus 1	DAB41737.1
Piscine myocarditis virus AL V-708	YP_004581250.1
Piscine myocarditis virus TT-2012	AEX97811.1
Pitaya virus X	YP_009046882.1
Point-Douro narna-like virus	YP_009388580.1
Potato mop-top virus	ALM54972.1
Potato virus X	BAB83122.1
Potato virus Y	BAI48919.1
Potexvirus sp.	YP_009553670.1

Psorophora varipes narnavirus	QBA55486.1
Pterostylis megabirnavirus-like	AOX47597.1
Pterostylis sanguinea totivirus A	AOX47551.1
Puccinia striiformis totivirus 3	ATO91011.1
Puccinia striiformis totivirus 4	ATO91013.1
Pythium nunn virus 1	YP_009551507.1
Pythium polare RNA virus 1	YP_009552275.1
Pythium splendens RNA virus 1	BBJ21453.1
Red clover powdery mildew-associated totivirus 4	BAT62484.1
Red clover powdery mildew-associated totivirus 7	YP_009182195.1
Red clover powdery mildew-associated totivirus 8	BAT62492.1
Red clover powdery mildew-associated totivirus 9	YP_009182198.1
Rhizoctonia mitovirus 1 RS002	AHL25281.1
Rhizoctonia oryzae-sativae mitovirus 1	YP_009249807.1
Rhizoctonia solani bipartite-like virus 1	QDW81299.1
Rhizoctonia solani mitovirus 2	ALD89121.1
Rhizoctonia solani mitovirus 23	QDW65458.1
Rhizoctonia solani mitovirus 28	QDW65418.1
Rhizoctonia solani mitovirus 6	ALD89125.1
Rhizoctonia solani ourmia-like virus 1 RNA 1	ALD89131.1
Rhizophagus irregularis mito virus 2	AXY40443.1
Rhizophagus sp. RF1 mitovirus	YP_009552787.1
Rhizopus microsporus 20S narnavirus	QBC65280.1
Rhizopus microsporus 23S narnavirus	QBC65281.1
Rhododendron virus A	YP_003868436.1
Riboviria sp.	QDH87296.1
Riboviria sp.	QDH86884.1
Riboviria sp.	QDH87296.1
Riboviria sp.	QDH87038.1
Riboviria sp.	QDH88192.1
Riboviria sp.	QDH89666.1
Riboviria sp.	QDH89742.1
Riboviria sp.	QDH87071.1
Riboviria sp.	QDH87807.1
Riboviria sp.	QDH90244.1
Rice ragged stunt virus	NP_620541.1
Rosellinia necatrix quadrivirus 1	YP_005097975.1
Sacbrood virus	AIW58914.1
Saccharomyces 20S RNA narnavirus	NP_660178.1
Saccharomyces 23S RNA narnavirus	NP_660177.1
Saccharomyces cerevisiae virus L-A-28	AMV49327.1

Saccharomyces cerevisiae virus L-A-28	AMV49327.1
Saccharomyces kudriavzevii virus L-A1	YP_009328931.1
Saccharomyces paradoxus virus L-A-21	ATL63180.1
Saccharomyces paradoxus virus L-A-66	AYN80723.1
Sanxia water strider virus 20	YP_009336893.1
Scheffersomyces segobiensis virus L	YP_009507829.1
Schlumbergera virus X	AJF19167.1
Schlumbergera virus X	YP_002341559.1
Sclerotinia homoeocarpa mitovirus	AAO21337.1
Sclerotinia nivalis victorivirus 1	YP_009259368.1
Sclerotinia sclerotiorum botybirnavirus 3	QDF82045.1
Sclerotinia sclerotiorum deltaflexivirus 1	AMD16208.1
Sclerotinia sclerotiorum double-stranded RNA virus 3	AND83002.1
Sclerotinia sclerotiorum mitovirus 1 HC025-A	AWY10963.1
Sclerotinia sclerotiorum mitovirus 12	AHF48628.1
Sclerotinia sclerotiorum mitovirus 14-A	AWY10977.1
Sclerotinia sclerotiorum mitovirus 2	AGC24231.1
Sclerotinia sclerotiorum mitovirus 2-A	AWY10964.1
Sclerotinia sclerotiorum mitovirus 3	YP_009182164.1
Sclerotinia sclerotiorum mitovirus 6-A	AWY10967.1
Sclerotinia sclerotiorum mitovirus 7-A	AWY10969.1
Sclerotinia sclerotiorum mitovirus 7-A2	AWY10970.1
Sclerotinia sclerotiorum mycoreovirus 4	YP_009252403.1
Sclerotinia sclerotiorum mycotymovirus 1	QDF82047.1
Sclerotinia sclerotiorum tymo-like RNA virus 4	AWY10995.1
Sclerotinia sclerotiorum umbra-like virus 1	YP_009253998.1
Sclerotium hydrophilum virus 1	YP_009273017.1
Sclerotium rolfsii unassigned dsRNA virus 1	AZF86115.1
Sclerotium rolfsii unassigned dsRNA virus 2	AZF86116.1
Setosphaeria turcica mitovirus 1	AZT88625.1
Shahe tombus-like virus 1	YP_009336903.1
Sherlock virus	QED21500.1
Shuangao chryso-like virus 1	ASA47445.1
Shuangao toti-like virus	YP_009336732.1
Solenopsis midden virus	QBL75907.1
Southern tomato virus	BAX64141.1
Soybean leaf-associated mitovirus 3	ALM62243.1
Soybean leaf-associated ourmiavirus 1	YP_009666497.1
Sphaeropsis sapinea RNA virus 2	NP_047560.1
Spissistilus festinus virus 1	YP_003800001.1
Strawberry mild yellow edge virus	AKN20462.1

Thelebolus microsporus totivirus 1	AZT88643.1
Tobacco mild green mosaic virus	ABH10605.1
Tobacco rattle virus	NP_620669.1
Tolypocladium ophioglossoides totivirus 1	AZT88645.1
Tomato brown rugose fruit virus	YP_009182168.1
Tomato bushy stunt virus	NP_062897.1
Tomato mottle mosaic virus	ALG02427.1
Totiviridae sp.	DH91165.1
Trichoderma harzianum bipartite mycovirus 1	YP_009553330.1
Tulip breaking virus	AHI04506.1
Tulip virus X	AYE20179.1
Turnip yellow mosaic virus	NP_733819.1
Umbelopsis ramanniana virus 2	VFI65724.1
Umbelopsis ramanniana virus 4	VUD77425.1
Ustilaginoidea virens RNA virus 5	YP_009182167.1
Ustilaginoidea virens RNA virus M	YP_009094186.1
Valsa cypri partitivirus	AIS37548.1
Verticillium albo-atrum partitivirus-1	AIE47664.1
Vespa velutina partiti-like virus 2	ATY36110.1
Wallemia sebi mycovirus 1	ALO50138.1
Watercress white vein virus	AFC95826.1
Wenling narna-like virus 4	YP_009337133.1
Wenling narna-like virus 5	YP_009337146.1
Wenling narna-like virus 6	APG77272.1
Wenling narna-like virus 7	YP_009337166.1
Wenling narna-like virus 8	APG77263.1
Wenling tombus-like virus 1	YP_009337158.1
Wenling tombus-like virus 1	YP_009337158.1
Wenzhou crab virus 5	YP_009337111.1
White clover cryptic virus 1	YP_086754.1
Wilkie narna-like virus 1	ASA47364.1
Wuhan insect virus 18	YP_009342440.1
Wuhan insect virus 26	YP_009342428.1
Wuhan insect virus 27	YP_009342434.1
Xanthophyllomyces dendrorhous virus L1B	YP_009507835.1
Xysticus cristatus iflavirus	APD13905.1
Zea mays chrysovirus 1	AYD75753.1
Zostera marina amalgavirus 1	YP_009362302.1
Zygocactus virus X	AFI57885.1
Coxsackievirus B3	K02709.1
Sclerotinia sclerotiorum debilitation-associated virus	Q6YI57.2

Aspergillus foetidus slow virus 1	YP_009508249.1
Beauveria bassiana victorivirus 1	YP_009508251.1
Rosellinia necatrix victorivirus 1	YP_008130308.1
Helminthosporium victoriae virus 190S	NP_619670.2
Leishmania RNA virus 1	APT68189.1
Leishmania RNA virus 2	AHK06416.1
Totivirus Tuber aestivum virus 1	YP_009507833.1
Saccharomyces cerevisiae virus L-A	NP_620495.1
Xanthophyllomyces dendrorhous virus L1A	YP_007697651.1
Giardia lamblia virus	NP_620070.1
Ustilago maydis virus H1	NP_620728.1
Trichomonas vaginalis virus 1	YP_009162330.1
Slow bee paralysis virus	ACY25863.1
Rhizoctonia solani barnavirus 1	KP900904.2
Colobanthus quitensis associated barnavirus 1	MG686618.1
Mushroom bacilliform virus	U07551.1
Aedes pseudoscutellaris reovirus isolate France	Q2Y0E9.1
Zika virus	AVW85813.1
Trichomonas vaginalis virus 2	NP_624323.2
Cryphonectria parasitica mycoreovirus-1 (9B21)	Q7TDB6.1
Colorado tick fever virus Florio N-7180	Q9DSQ0.1

Chapter 3 Evidence for the cospeciation of mycoviruses and their hosts

Abstract

Though historically thought to be the dominant mode amongst hosts and symbionts, cospeciation has since been found to be quite rare. Amongst RNA viruses, the alternative, host-switching, is especially common. RNA viruses in fungi, however, appear to be an exception because they lack an external transmission route. Here, we revisit analyses of mycovirus-host coevolution to include viruses from lineages of hosts previously unsampled, including the early-diverging lineages Neocallimastigomycota, Blastocladiomycota, Chytridiomycota, Mucoromycota, and Zoopagomycota. Because challenges associated with RNA virus sequencing have partially been responsible for the lack of mycovirus data from these diverse fungal lineages, we compared three NGS sequencing methods. Our results corroborate previous findings of cospeciation between families of mycoviruses and their fungal hosts, suggesting that a common fungal ancestor was host to multiple families of viruses that have cospeciated with their hosts. Statistical tests of some individual host-virus associations do not support this, however, which suggests that some host-switching has occurred despite the overall pattern of cospeciation. Our results have implications for our understanding of mycovirus-host interactions, mycovirus ecology, and the potential application of mycoviruses for biological control of fungal pathogens.

Introduction

Host-switching of a symbiont from an historic host species to a naive host species is a phenomenon often held responsible for disease outbreaks. Thus, the degree of host-specificity of a symbiont is a key component of disease dynamics, and factors contributing to host-specificity

are central to understanding disease ecology more generally. Research on the mechanism and frequency of cross-species transmission is essential to understand factors contributing to disease emergence and control and has increasingly received attention for the implications to public health and agriculture.

It has historically been suggested that hosts and symbionts diversify via “cospeciation”. The underlying assumption is that host-symbiont interactions are obligate or highly specialized and so speciation processes in the host result in isolation and subsequent speciation of the symbiont. This idea was described in the early twentieth century by Fahrenholz, and the idea that “parasite phylogeny mirrors that of its host” came to be known as Fahrenholz rule (Fahrenholz 1913). Alternatively, symbionts can undergo “host-shift speciation”, in which the symbiont specializes on a different host than its immediate ancestor (reviewed in De Vienne 2013).

Methods to distinguish between these processes involve tests of phylogenetic congruence of host and symbiont and fall into two main classes. Event-based methods rely on topology only and compare the observed number of cospeciation events with cospeciation events generated by random association of hosts and symbionts. These methods effectively test topological congruence but can overestimate the occurrence of cospeciation since alternative processes other than cospeciation can also result in phylogenetic congruence. For instance, host-switching of a symbiont to a closely related host species also results in apparent cophylogeny. Distance-based methods can overcome this to some extent, provided that input trees are accurate, because these methods consider both branch ordering and branch lengths (i.e. time of speciation events). Processes such as independent speciation of the symbiont, lack of speciation of symbiont with host speciation, and symbiont extinction in some host lineages further complicate phylogenies and can reduce the signal of cophylogeny, even when cospeciation is occuring (Page 2003).

Thus, while there are clear expectations for host and symbiont phylogenetic topologies under cospeciation and host-shift speciation, the practical challenge of discerning between them is nontrivial.

Studies using robust methods have demonstrated that cospeciation is, perhaps surprisingly, rare. De Vienne *et al*. summarize studies demonstrating some level of cospeciation; given the substantial interest in this topic over the past century, it is remarkable that this was possible in just a few pages. Indeed, the authors remark that "convincing cospeciation between host and symbiont trees is seldom found except for a few mutualist associations, most often involving vertically transmitted symbionts."

Viruses of fungi, termed mycoviruses, are a particularly interesting test case for cospeciation research. They are propagated through vertical transmission in spores as well as, in a presumably more limited manner, through horizontal transmission to compatible members of the same species via hyphal fusion. There are currently no known mycoviruses that lyse their host cell and mycoviral infections are consistently described as "persistent" and "latent". Some are known mutualists (Márquez 2007; Schmitt & Breinig 2002). Given this biology, one hypothesis of mycoviral origins is infection of a common ancestor and codivergence of viruses and Fungi (Son 2015). Previous work has demonstrated some evidence of cospeciation in some mycoviral families, including Totiviridae and Partitiviridae, as well as a lack of cospeciation in others, including Chrysoviridae and Narnaviridae (Göker 2011).

Nearly a decade later, the question of mycovirus-host cospeciation is worth revisiting.

Importantly, at the time of Göker and colleagues' work, mycoviruses were known and sequenced from only two fungal phyla (Ascomycota and Basidiomycota), of which there are at least 8 recognized (Spatafora 2016; James 2020). The authors thus acknowledged the limitation

in their findings due to the lack of sampling. The inclusion of early-diverging lineages allows a test of how deep in evolutionary time viruses and fungi have coevolved, given the 1-billion-year age of the kingdom (Berbee 2017). Contributing to the lack of sampling are challenges associated with viral sequencing. While classic methods for screening fungi for RNA viruses are relatively simple and fast (Morris & Dodds 1979; Okada 2015), sequencing these viruses is more burdensome. Known mycoviruses are predominantly composed of double-stranded RNA (dsRNA) genomes, often with complex secondary structure, which has posed a significant challenge for genomic sequencing. dsRNA is the most stable nucleic acid, and so melting the strands and keeping them separated at temperatures amenable to polymerases is nontrivial. Further, for viruses for which the sequence is unknown, cDNA must be generated by random priming which, given the complex secondary structures and stability of dsRNA, often results in discontinuous stretches of cDNA and thus downstream assembly challenges.

Here we revisit the "ancient coevolution hypothesis" of mycovirus origin with robust testing that incorporates data across the fungal kingdom, including newly sequenced viruses found in the Neocallimastigomycota, Chytridiomycota, Blastocladiomycota, Mucoromycota, and Zoopagomycota. To assess the ability of various methods to accommodate the challenges of sequencing dsRNA, we compare three alternative methods commonly used in mycoviral research.

Methods

Fungal cultivation, mycovirus screening, sequencing, and mycovirus assembly

Methods of fungal cultivation, fungal screening for mycoviruses, dsRNA purification, reverse-transcription, sequencing of cDNA by PacBio and Illumina MiSeq, sequencing of mRNA by Illumina MiSeq, and mycovirus assembly were described in Chapter 2. Methods of

sequencing fungal ribosomal regions by Oxford Nanopore Technology were previously described (Simmons 2020). Additional mycovirus RdRp and fungal 28S sequences were retrieved from GenBank (Supplementary Table 3.1). Where host sequences were unavailable, the sequence of a close relative of the same genus was used only when no other viruses of that genus were present; otherwise, the virus and host were removed from the analysis.

Phylogenetic analyses

We aligned sequences with MAFFT version 7 using the E-INS-i algorithm (Katoh 2013) and selected conserved sites with Gblocks (Talavera and Castresana 2007) implemented in Seaview v4 (Guoy 2010) employing options for a less stringent selection. We determined the best-fit model of amino acid substitution for viral alignments with ProtTest 3.4 (Darriba 2017) and reconstructed trees with the maximum likelihood approach implemented in RaxML by the rapid bootstrap analysis (-f a) with 100 bootstrap pseudoreplicates (Stamatakis 2014).

Tests of cospeciation

The program Parafit tests the null hypothesis that two input phylogenies associate at random with respect to one another, the evolutionary interpretation being that the two trees evolved independently. Operationally, the program computes principle coordinates for the two distance matrices and takes a fourth-corner approach (Legendre 1997) to estimate the parameters that cross the principle coordinates of one phylogeny with the other. The sum of squares of the resultant fourth-corner matrix are computed in the ParafitGlobal statistic. The tests of individual linkages between phylogenies function by the idea that if a relationship is removed from the analyses, the ParafitGlobal statistic should decrease if the relationship was important to the overal global fit of the cophylogeny (Legendre 2002).

We evaluated the cophylogeny of mycoviruses and their hosts using Parafit (Legendre 2002) implemented by the ape package (version 5.4) (Paradis and Schliep 2019) in R, using patristic distances from the maximum likelihood phylogenies, 999 permutations, both the global fit and individual links tests, and both the "cailliez" and "lingoes" corrections for negative eigenvalues. The two correction methods were similar for all tests, so the "cailliez" values are reported. We drew tanglegrams with TreeMap 3.0 (Charleston 2011), using the "untangle" function to minimize the number of intersections in host-virus links shown. We used SankeyMATIC (sankeymatic.com) to create the sankey diagram.

Alignment, code and input data for all analyses are available at github.com/jimyers.

Results

Comparison of sequencing approaches

We compared different approaches by sequencing the same isolates by multiple methods (Figure 3.1). Common viruses were recovered by each method but, unsurprisingly, they were represented by contigs varying in length. For example, a 1436 bp RNA-dependent RNA polymerase (RdRp) contig was produced by PacBio sequencing of *Zopfochytrium polystomum* WB288 dsRNA while the same isolate sequenced by Illumina MiSeq mRNA sequencing resulted in only 328bp of identical partial sequence. PacBio did not always outperform, however; by mRNA MiSeq a 3773 bp contig from *Allomyces* sp. JMM01 was produced while PacBio resulted in two contigs of 1921 bp and 1703 bp identical to either the 5' or 3' end of the MiSeq-generated contig, respectively, leaving a 134 bp gap in the RdRp domain. Similarly, comparison of Illumina MiSeq of purified dsRNA to PacBio was inconclusive. The dsRNA of a strain of the Basidiomycete *Ustilago maydis* harboring the H1 "killer" virus was sequenced by PacBio and a contig was assembled that is nearly-complete and nearly-identical to the GenBank sequence

(6036 bp/6099 bp; 93% identity), whereas Illumina MiSeq of dsRNA cDNA produced two short contigs of 363 bp and 889 bp. However, a previously unreported mitovirus was also recovered by each method, this time with dsRNA MiSeq outperforming PacBio in terms of contig length (2434 bp and 592 bp, respectively). These discrepancies could be caused by inherent variation in any number of factors including random priming, secondary structure, affinity of different reverse-transcription enzymes, and the RNA quality of different extraction preparations. Based on these mixed results, we are reluctant to suggest any one method for de novo sequencing over another. Instead, more representative sequences might be achieved through the sequencing of libraries prepared as pools of the cDNA of many independent dsRNA replicates—thus diversifying the library from the biases of any single RNA extraction or single reverse-transcription reaction by random priming.

Interestingly, additional viral contigs were often uncovered by the different methods (Figure 3.1). In addition to the factors mentioned above, it is possible that some methods are better able to handle certain properties of some viruses (such as particularly complex secondary structures) compared to other methods. The differences could also be a result of stochastic differences in population sizes of the different viruses at various time points in a given culture. If so, it will be important to keep in mind that, even through sequencing, we may only be able to reveal the virome of an organism as a snapshot in time.

Tests of cospeciation

Parafit tests the null hypothesis that host and parasite phylogenies are independent (i.e. one phylogeny is random with respect to the other), which was rejected for all four viral families tested (ParafitGlobal $p < 0.001$; Table 3.1), suggesting that cospeciation has occurred. Interestingly, however, viruses in the basal fungal lineages varied in whether or not they

contributed to the overall global cophylogeny. For instance, Totiviridae viruses from Blastocladiomycota had significant individual links statistics ($p < 0.05$), indicating that these virus-host associations are unlikely to be random with respect to the coevolutionary structure. However, in this same viral family, viruses from Zoopagomycota and Mucoromycota hosts had insignificant individual links statistics and so the null hypothesis that these associations are random cannot be rejected (Supplementary Table 3.2). Thus, while cospeciation remains the dominant signal overall, there is some evidence of deep rearrangement (Figures 3.2-3.5). Previous studies have shown that with more viral sampling, the greater the likelihood of finding host-switching events (Geoghegan 2017). Our data perhaps hint that with even greater representation of early-diverging fungi, an alternative global coevolutionary structure could be revealed with less signal of cospeciation.

Global patterns of mycovirus-host evolutionary relationships

The four viral families tested have phylogenetically widespread representation across the funal kingdom, but with some meaningful variation (Figure 3.6). It has recently been suggested that mitoviruses have a common origin in the ancestor of Mucoromycota and Zoopagomycota (Nibert 2019). Consistent with this hypothesis, Mitoviridae appear to be absent in the Blastocladiomycota, Chytridiomycota, and Neocallimastigomycota and there is overall statistical support for cospeciation (Table 3.1). Chrysoviridae are represented in 4 of the 8 phyla, including two early-diverging lineages (Mucoromycota and Blastocladiomycota). This viral family had the smallest sample size and it is likey that with futher sampling there will be greater fungal phylogenetic representation. Partitiviridae are represented in 5 of the 8 phyla, which, interestingly, includes one of the most basal clades, Neocallimastigomycota. The Totiviridae are

the most phylogenetically widespread appearing in 6 of the 8 phyla, but not yet idenified in the Glomeromycota or the Neocallimastigomycota.

Importantly, Figure 3.6 was drawn according to the number of viruses in each group, which is not necessarily equivalent to the number of hosts (i.e. numerous viruses have been identified in some individual host species). Accordingly, the width of the bars in the middle correspond to the number of viruses in each fungal class, rather than the number of fungi represented from that class. This provides visualization of the taxonomy of the fungal "source" of each virus, but consequentially denies visualization of the true virus:host ratio which can be greater than 1:1. For instance, 85 mitoviruses included in the analysis had only 41 hosts (Table 3.1). This suggests that Mitoviridae also undergos speciation via duplication, as evidenced by the 17 mitoviruses in *Sclerotinia sclerotiorum* alone, which is not even exhaustive of all that have been identified in this species (Figure 3.4).

Discussion and Conclusions

Technical challenges due to the common nucleic acid type of mycovirus have been prohibitive of mycoviral characterization, contributing to the overall lack of viral sampling in fungi and consequently limiting the power of evolutionary tests. We compared three methods of dsRNA sequencing. Each method has strengths and weaknesses, and no one method clearly outperforms the others. We suggest that, regardless of sequencing method used, outcomes may be improved by pooling multiple independent cDNA libraries.

Our study revisited tests of the origins of mycoviruses, increasing sample sizes and the breadth of fungal taxa sampled. Unlike a previous study, we also reconstructed host trees using sequence data, rather than using taxonomy alone to construct cladograms for cophylogenetic analyses. Thus, the phylogenetic distances of host and viruses were meaningful, and empowered

more robust testing of the coevolutionary structure. More viral data from fungi in early-diverging lineages is warranted but overall our findings thus far support the parsimonious hypothesis that mycoviral families share a common form of speciation, which is cospeciation with their hosts. This result corroborates findings for the largest families tested in Göker *et al*., Totiviridae, Partitiviridae, and Mitoviridae. Unlike the current study, however, those authors did not find statistically significant evidence of cospeciation in Chrysoviridae, likely due to low sample size. They found that sample size had a significant effect on the number of host-virus links that contributed to overall global fit, and, at the time, Chrysoviridae was much more poorly sampled than it is today ($n = 8$ in Göker 2011; $n = 32$ in present study).

Global statistics reject the null hypothesis of independent phylogeny between hosts and viruses for all four viral families tested ($p < 0.001$). However, many associations, including some between basal fungi and their viruses, do not significantly contribute to the overall global fit of the cophylogenies. Given the phylogenetic distance of the fungal lineages sampled, per cospeciation we would expect the viruses of these fungi to be similarly distant from other viruses, and thus be impactful to the signal of cophylogeny found. For example, under cospeciation the ancestors of a virus in an Ascomycete would have diverged from those of a related virus in a Chyridiomycete at a time coinciding with the divergence of the common fungal ancestor, approximately 500 million years ago. We would expect these viral lineages to have accumulated substantial genetic distance in that time, in their distinct hosts. Such a signal of great genetic distance between viruses in distantly related hosts would, logically, be powerful contributions to the overall global fit of the cophylogenies. Surprisingly, in every viral family tested, there were instances in which this was not true (Supplementary Tables 3.2-3.5). Thus, it is most probable that some host-switching has occurred.

Additionally, there are instances of coinfection in some species, which appears to have resulted from speciation of viruses via duplication within a host as evidenced by closely related but distinct viruses from the same fungal host. For instance, mitoviruses in *E. muscae* and *S. sclerotiorum* (Figure 3.4), partitiviruses in *U. virens* and *G. abietina* (Figure 3.3), chrysoviruses in *Allomyces* sp. (Figure 3.5), and totiviruses in *Cladochytrium* sp., *P. striliformis*, and *Saccharomyces* spp. (Figure 3.2). It is possible that some lineages of viruses may be more prone to this mechanism of speciation than others. Alternatively, it may be that some fungi are more prone to viruses that speciate via duplication. In Chapter 2, we identified massive coinfection in an isolate of *Cladochytrium sp*., in multiple species of *Allomyces*, and in *Kickxella alabastrina*. The viruses of these organisms appear to be closely related and are likely to have arisen via duplication. Many of these viruses did not resolve into the current viral taxonomy at the family level, however, and thus were not included in this analysis. Future studies involving additional fungal sampling should include viral taxonomic classification in order to include potentially novel viral clades in analyses. Further, future work should aim to quantify host switching relative to speciation by duplication, as they are not mutually exclusive.

Previous hypotheses of mycoviral origins have suggested a common ancestor of Fungi was infected with a virus and cospeciation ensued (Göker 2011; Bruenn 1993). Our data support this general hypothesis, albeit with occasional host-switches. Viruses recently identified in basal phyla Chytridiomycota and Blastocladiomycota appear to be in viral families Totiviridae, Chrysoviridae, and Partitiviridae (Figure 3.6). Thus, it appears that the process of cospeciation occurred for several families of viruses and their fungal hosts, which suggests that the common fungal ancestor was likely host to multiple virus types. It is important to articulate this nuance to the cospeciation hypothesis, lest it be misconstrued that these viral families evolved within

Fungi. Instead, most mycovirus lineages appear to be older than the age of the most recent common ancestor of fungi.

It is interesting to consider the implications of ancient mycoviral infection, and particularly their vertical transmission strategy. Extant mycovirus-host interactions are commonly asymptomatic, which is consistent with disease theory that vertical transmission selects for reduced virulence (Ebert 2013). However, a more ecologially holistic view would consider this not just through the lens of disease theory, but also from the symbiosis perspective. Mutualisms are also frequently maintained by vertical transmission, but viruses are under-appreciated as potential mutualists despite numerous examples of beneficial viruses (Xu 2008; Márquez 2015; Roossinck 2015; Márquez and Roossinck 2012; Roossinck 2011). Though we suggest that mycoviruses are more mutualistic than appreciated, interactions are likely complex and, perhaps, dependent on the environmental context.

Cospeciation is an uncommon phenomenon, and precedence suggests it primarily occurs in vertically transmitted mutualists (De Vienne *et al*. 2013). Our findings suggest that mycoviruses may be unique amongst RNA viruses, as a recent comprehensive study found host-switching to be the dominant mode of speciation in RNA viruses (Geoghegan 2017). Thus, our findings have implications to our understanding of mycovirus ecology and interactions with their hosts.

Figures

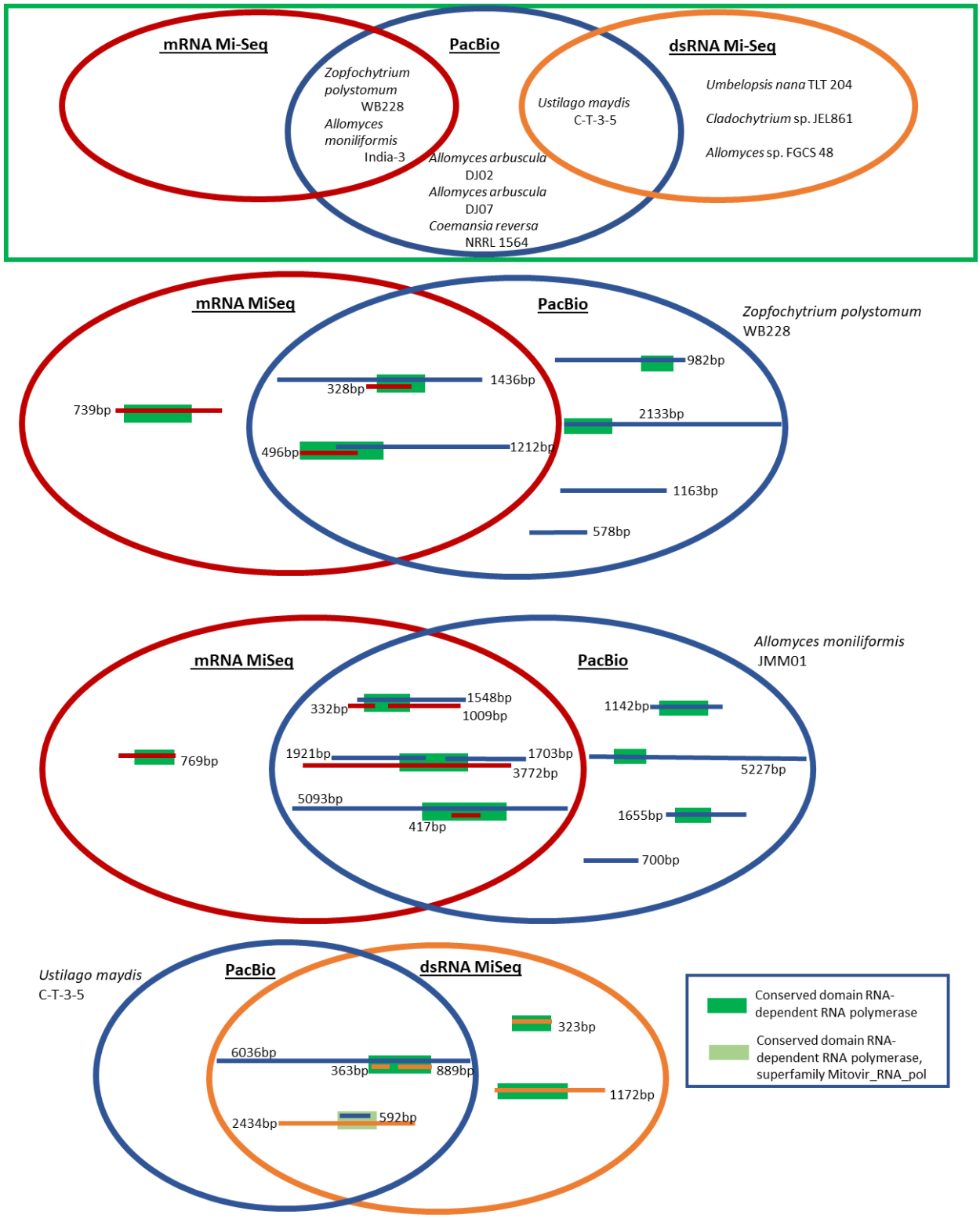

mRNA Mi-Seq
PacBio
dsRNA Mi-Seq
Zopfochytrium polystomum WB228
Allomyces moniliformis India-3
Allomyces arbuscula DJ02
Allomyces arbuscula DJ07
Coemansia reversa NRRL 1564
Ustilago maydis C-T-3-5
Umbelopsis nana TLT 204
Cladochytrium sp. JEL861
Allomyces sp. FGCS 48
mRNA MiSeq
PacBio
Zopfochytrium polystomum WB228
739bp
328bp
1436bp
496bp
1212bp
982bp
2133bp
1163bp
578bp
mRNA MiSeq
PacBio
Allomyces moniliformis JMM01
769bp
332bp
1548bp
1009bp
1921bp
1703bp
3772bp
5093bp
417bp
1142bp
5227bp
1655bp
700bp
Ustilago maydis C-T-3-5
PacBio
dsRNA MiSeq
6036bp
363bp
889bp
2434bp
592bp
323bp
1172bp
Conserved domain RNA-dependent RNA polymerase
Conserved domain RNA-dependent RNA polymerase, superfamily Mitovir_RNA_pol

Figure 3.1 Results of sequencing the viromes of multiple fungi by three different NGS methods are compared in a venn diagram-like format. Three representative viromes are shown (*Zopfochytrium polystomum* WB228, *Allomyces sp. JMM01*, and *Ustilago maydis* C-T-3-5). Lines indicate viral genome, colored according to sequencing method, with green boxes denoting approximate relative location of RNA-dependent RNA polymerase (RdRp) domains. Assembled contig sizes are indicated.

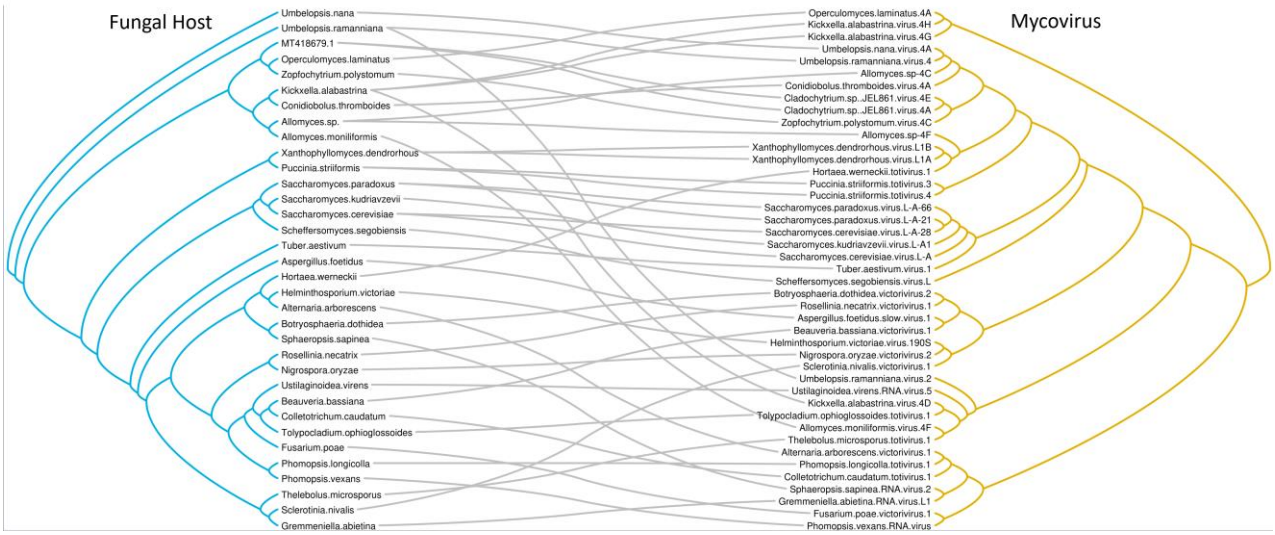


Figure 3.2 Cophylogeny of Totiviridae and their fungal hosts. Overall, there is a clear signal of congruence, evidenced by parallel connections between viruses and hosts. Some connections appear to violate this, however, including those between *A. moniliformis*, *H. werneckii*, and *K. alabastrina* and their respective viruses.

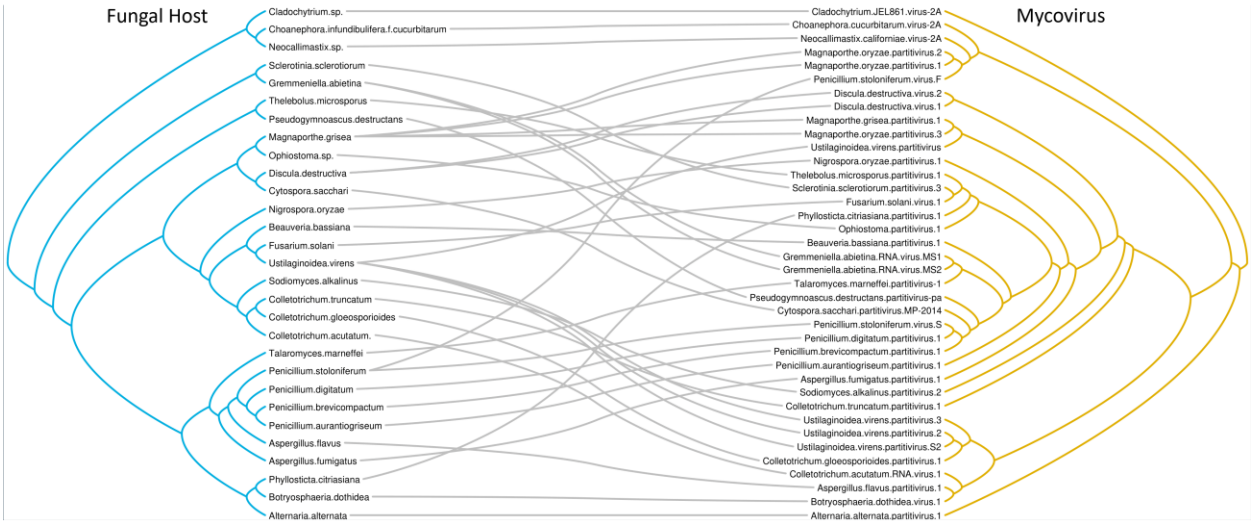


Figure 3.3 Cophylogeny of Partitiviridae and their fungal hosts. There is a signal of congruence, evidenced by parallel connections between viruses and hosts, numerous connections appear to defy this.

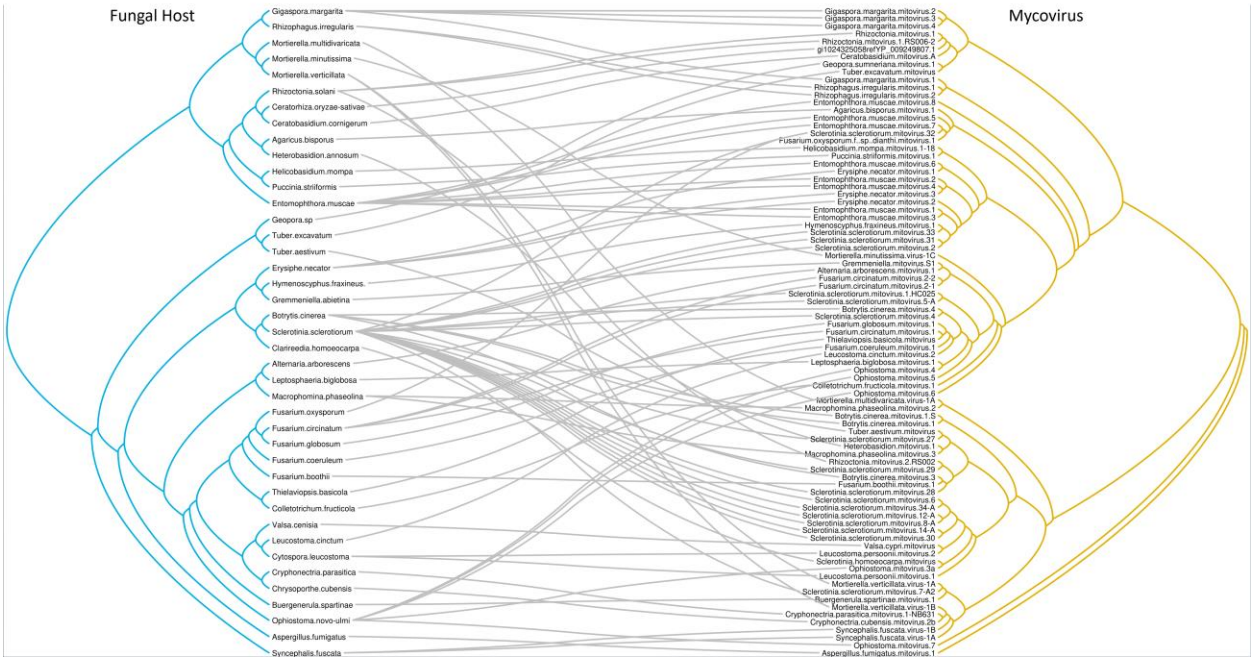


Figure 3.4 Cophylogeny of Mitoviridae and their fungal hosts. Moreso than others, this cophylogeny has numerous instances of apparent viral speciation by duplication, evidenced by the multitude of closely related viruses in one host (ex. *S. sclerotiorum*, *E. muscae*, *G. margarita*, *B. cinerea*, etc.)

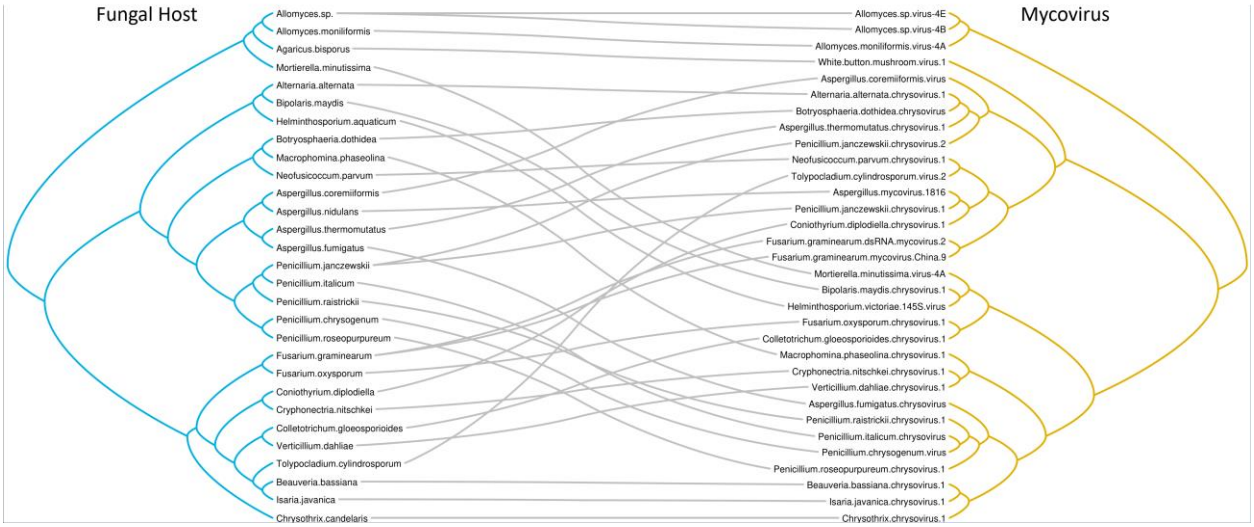


Figure 3.5 Cophylogeny of Chrysoviridae and their fungal hosts shows primarily phylogenetic congruence and some possible instances of incongruence.

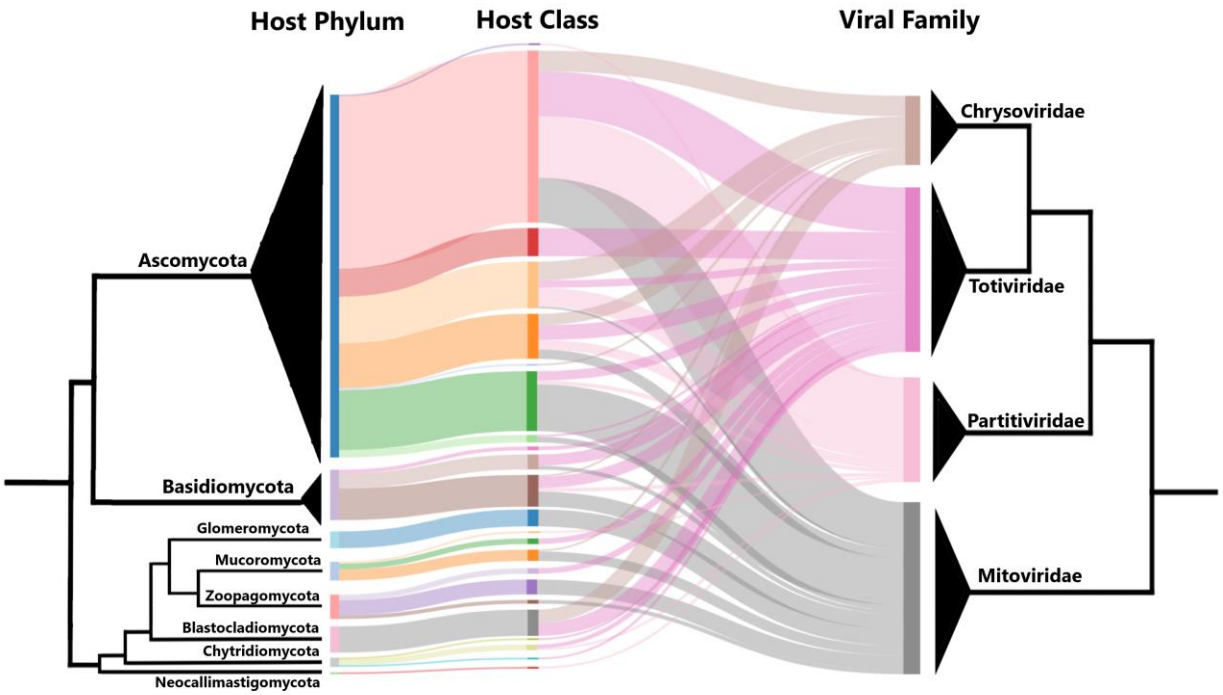


Figure 3.6 Sankey diagram displaying data for all mycoviruses used in the analyses. On the left is phylum to which viral hosts belong, which feeds into host class, and then, on far right, viral family. All four viral families are hosted by diverse fungi in multiple taxonomic groups. Of the four, Totiviridae has the highest representation across fungal phyla.

Tables

Table 3.1 Results of parafit global tests of cospeciation. For all four viral families tests, the null hypothesis of random association between virus and host was rejected ($p < 0.05$).

	Totiviridae	**Partitiviridae**	**Mitoviridae**	**Chrysoviridae**
N (fungus)	34	29	41	29
N (virus)	43	40	85	32
ParafitGlobal (nperm =9999; correction= "cailliez")	641.7488	273.3006	840.662	144.2948
p-value	1.00E-04	6.00E-04	1.00E-04	0.005
ParafitGlobal (nperm =9999; correction = "lingoes")	446.2182	176.0191	474.8463	104.6771
p-value	2.00E-04	0.0029	1.00E-04	0.006

References

Berbee, ML, TY James, and C Strullu-Derrien. 2017. Early diverging fungi: diversity and impact at the dawn of terrestrial life. *Annual Review of Microbiology* **71** (1): 41–60. https://doi.org/10.1146/annurev-micro-030117-020324.

Bonfante, P, and F Venice. 2020. Mucoromycota: going to the roots of plant-interacting fungi. *Fungal Biology Reviews* **34** (2): 100–113. https://doi.org/10.1016/j.fbr.2019.12.003.

Bruenn, JA. 1993. A closely related group of RNA-dependent RNA polymerases from double stranded RNA viruses. *Nucleic Acids Research* **21** (24): 5667–69. https://doi.org/10.1093/nar/21.24.5667.

Charleston M. 2011. Treemap 3.0 https://sites.google.com/site/cophylogeny/home

Darriba, D, GL Taboada, R Doallo, and D Posada. 2017. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27** (8): 1164–65. https://doi.org/10.1093/bioinformatics/btr088.ProtTest.

de Vienne, DM, G Refrégier, M López-Villavicencio, A Tellier, ME Hood, and T Giraud. 2013. Cospeciation vs host-shift speciation: methods for testing, evidence from natural associations and relation to coevolution. *New Phytologist* **198** (2): 347–85. https://doi.org/10.1111/nph.12150.

Ebert, D. 2013. The epidemiology and evolution of symbionts with mixed-mode transmission. *Annu. Rev. Ecol. Syst.* **44**, 623–43.

Fahrenholz H. 1913. Ectoparasiten und abstammungslehre. Zoologischer Anzeiger **41**, 371–374.

Freeman, KR, AP Martin, D Karki, RC Lynch, MS Mitter, AF Meyer, JE Longcore, DR Simmons, and SK Schmidt. 2009. Evidence that chytrids dominate fungal communities in high-elevation soils. *Proceedings of the National Academy of Sciences of the United States of America* **106** (43): 18315–20. https://doi.org/10.1073/pnas.0907303106.

Geoghegan, JL, S Duchêne, and EC Holmes. 2017. Comparative analysis estimates the relative frequencies of co-divergence and cross-species transmission within viral families. *PLoS Pathogens* **13** (2): 1–17. https://doi.org/10.1371/journal.ppat.1006215.

Ghabrial, SA, JP Castón, , D Jiang, ML Nibert, and N Suzuki. 2015. 50-Plus years of fungal viruses. *Virology*, 479–480, 356–368. https://doi.org/10.1016/j.virol.2015.02.034

Göker, M, C Scheuner, HP Klenk, JB Stielow, and W Menzel. 2011. Codivergence of mycoviruses with their hosts. *PLoS ONE* **6** https://doi.org/10.1371/journal.pone.0022252.

Gouy, M, S Guindon, and O Gascuel. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution* **27**(2), 221–224, https://doi.org/10.1093/molbev/msp259

James, T. Y., J. E. Stajich, C. T. Hittinger, and A. Rokas. (2020). Towards a fully resolved fungal tree of life. *Annu. Rev. Microbiol*. **74**: 291-313.

Katoh, K, and DM Standley. 2013. MAFFT: multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30** (4): 772–80. https://doi.org/10.1093/molbev/mst010.

Khalil, JYB, J Andreani, and B La Scola. 2016. Updating strategies for isolating and discovering giant viruses. *Current Opinion in Microbiology* **31**: 80–87. https://doi.org/10.1016/j.mib.2016.03.004.

Legendre, PR, R Galzin, and ML Harmelin-Vivien. 1997. Relating behavior to habitat: solutions to the fourth-corner problem. *Ecology* **78**,547–562.

Legendre, P, Y Desdevises, and E Bazin. 2002. A statistical test for host-parasite coevolution. *Systematic Biology* **51** (2): 217–34. https://doi.org/10.1080/10635150252899734.

Márquez, LM, and MJ Roossinck. 2012. Do persistent RNA viruses fit the trade-off hypothesis of virulence evolution? *Current Opinion in Virology* **2** (5): 556–60. https://doi.org/10.1016/j.coviro.2012.06.010.

Márquez, LM, RS Redman, RJ Rodriguez, and MJ Roossinck. 2007. A virus in a fungus in a plant : three-way symbiosis required for thermal tolerance. *Science* **315** (5811): 513–15.

Morris, T. J., & Dodds, J. A. (1979). Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. *Phytopathology*, 69(8), 854–858.

Nibert, M. L., Debat, H. J., Manny, A. R., Grigoriev, I. V., & De Fine Licht, H. H. (2019). Mitovirus and mitochondrial coding sequences from basal fungus *Entomophthora muscae*. *Viruses*, *11*(4), 6–8. https://doi.org/10.3390/v11040351

Okada, R, E Kiyota, H Moriyama, T Fukuhara, and T Natsuaki. 2015. A simple and rapid method to purify viral dsRNA from plant and fungal tissue. *Journal of General Plant Pathology* **81** (2): 103–7. https://doi.org/10.1007/s10327-014-0575-6.

Page, R. 2003. Tangled trees. Phylogeny, cospeciation, and coevolution. *The University of Chicago Press*.

Paradis E, and K Schliep. 2019. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, **35**, 526-528.

Roossinck, MJ. 2011. The good viruses:viral mutualistic symbiosis. *Nature Reviews Microbiology.* **9,** 99-108.

Roossinck, MJ. 2015. Move over, bacteria! Viruses make their mark as mutualistic microbial symbionts. *Journal of Virology* **89** (13): 6532–35. https://doi.org/10.1128/jvi.02974-14.

Schmitt, MJ, and F Breinig. 2002. The viral killer system in yeast: from molecular biology to application. *FEMS Microbiology Reviews* **26** (3): 257–76. https://doi.org/10.1016/S0168-6445(02)00099-2.

Simmons DR, AE Bonds, BT Castillo, RA Clemons, AD Glasco, JM Myers, N Thapa, PM Letcher, MJ Powell, JE Longcore, & TY James. 2020. The collection of zoosporic eufungi at the University of Michigan (CZEUM): introducing a new repository of fully barcoded Chytridiomyceta and Blastocladiomycota cultures. *IMA Fungus.* In press.

Son, M, J Yu, and KH Kim. 2015. Five questions about mycoviruses. *PLoS Pathogens* **11** (11): 5–11. https://doi.org/10.1371/journal.ppat.1005172.

Spatafora, JW, Y Chang, GL Benny, K Lazarus, ME Smith, ML Berbee, G Bonito, *et al*. 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* **108** (5): 1028–46. https://doi.org/10.3852/16-042.

Stamatakis, A. 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30** (9): 1312–13. https://doi.org/10.1093/bioinformatics/btu033.

Talavera, G, and J Castresana. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Systematic Biology* **56**, 564-577.

Xu, P, F Chen, JP Mannas, T Feldman, LW Sumner, and MJ Roossinck. 2008. Virus infection improves drought tolerance. *New Phytologist* **180** (4): 911–21. https://doi.org/10.1111/j.1469-8137.2008.02627.x.

Supplement

Supplemental Table 3.1 GenBank accessions of fungal host 28S genes.

Fungal Host	**Accession**
Agaricus bisporus	MH870798.1
Alternaria alternata	DQ678082.1
Alternaria arborescens	NG_069124.1
Aspergillus coremiiformis	NG_064111.1
Aspergillus flavus	AF109341.1
Aspergillus foetidus	MH866147.1
Aspergillus fumigatus	AF109331.1
Aspergillus nidulans	MH858232.1
Aspergillus thermomutatus	NG_055722.1
Beauveria bassiana	AF280637.1
Bipolaris maydis	MH875909.1
Botryosphaeria dothidea	DQ377849.1
Botrytis cinerea	KP671724.1
Buergenerula spartinae	DQ341492.1
Ceratobasidium cornigerum	AY152405.1
Ceratorhiza oryzae-sativae	MH873047.1
Choanephora infundibulifera f. cucurbitarum	AB536741.1
Chrysoporthe cubensis	JN940856.1
Chrysothrix candelaris	KF707640.1
Cladochytrium sp.	MT418679.1
Clarireedia homoeocarpa	MH867420.1
Colletotrichum acutatum	AF275542.1
Colletotrichum caudatum	Z18985.1
Colletotrichum fructicola	JN940418.1
Colletotrichum gloeosporioides	AY705727.1
Colletotrichum truncatum	KY347817.1
Conidiobolus thromboides	NG_058824.1
Coniothyrium diplodiella	AY339286.1
Cryphonectria nitschkei	AF408341.1

Cryphonectria parasitica	JN940858.1
Cytospora leucostoma	MK673086.1
Cytospora sacchari	MH866845.1
Discula destructiva	AF408359.1
Entomophthora muscae	DQ481224.2
Erysiphe necator	MT182949.1
Fusarium boothii	MH860690.1
Fusarium circinatum	MH874260.1
Fusarium coeruleum	MH878483.1
Fusarium globosum	U61661.1
Fusarium graminearum	HQ147601.1
Fusarium oxysporum	KU729130.1
Fusarium poae	JN938907.1
Fusarium solani	DQ236726.1
Geopora sp.	KC012679.1
Gigaspora margarita	JF817011.1
Gremmeniella abietina	FN868861.1
Helicobasidium mompa	NG_059418.1
Helminthosporium victoriae	NG_059656.1
Heterobasidion annosum	KJ651519.1
Hortaea werneckii	NG_057773.1
Hymenoscyphus fraxineus	HM145907.1
Isaria javanica	NG_059048.1
Kickxella alabastrina	KF848900.1
Leptosphaeria biglobosa	KT389759.1
Leucostoma cinctum	AF408366.1
Macrophomina phaseolina	DQ377912.1
Magnaporthe grisea	KP144449.1
Mortierella minutissima	NG_042557.1
Mortierella multidivaricata	MH872890.1
Mortierella verticillata	KC018446.1
Neocallimastix sp.	MK398240.1
Neofusicoccum parvum	NG_042409.1

Nigrospora oryzae	FJ176892.1
Ophiostoma novo-ulmi	MH874409.1
Penicillium aurantiogriseum	AF003355.1
Penicillium brevicompactum	MH877526.1
Penicillium chrysogenum	FJ890400.1
Penicillium digitatum	MH874465.1
Penicillium italicum	NG_069652.1
Penicillium janczewskii	NG_069617.1
Penicillium raistrickii	KF880952.1
Penicillium roseopurpureum	NG_064128.1
Penicillium stoloniferum	MH875697.1
Phomopsis longicolla	FJ755236.1
Phomopsis vexans	AB104644.1
Phyllosticta citriasiana	KF766379.1
Pseudogymnoascus destructans	KJ938427.1
Puccinia striiformis	DQ417403.1
Rhizoctonia solani	JX576188.1
Rhizophagus irregularis	LR731874.1
Rosellinia necatrix	AY083824.1
Saccharomyces cerevisiae	NG_042623.1
Saccharomyces kudriavzevii	NG_042469.1
Saccharomyces paradoxus	NG_055028.1
Scheffersomyces segobiensis	U45742.1
Sclerotinia nivalis	KM211700.1
Sclerotinia sclerotiorum	MN078807.1
Sodiomyces alkalinus	MH878336.1
Sphaeropsis sapinea	DQ377893.1
Syncephalis fuscata	KY001786.1
Talaromyces marneffei	MH874009.1
Thelebolus microsporus	LC514937.1
Thielaviopsis basicola	KM495397.1
Tolypocladium cylindrosporum	NG_067391.1
Tolypocladium ophioglossoides	JN941405.1

Tuber aestivum	KF523368.1
Tuber excavatum	FJ809825.1
Umbelopsis nana	NG_058036.1
Umbelopsis ramanniana	AF113463.1
Ustilaginoidea virens	AF245299.1
Valsa cenisia	AF408385.1
Verticillium dahliae	NG_069484.1
Xanthophyllomyces dendrorhous	AF189871.2
Zopfochytrium polystomum	MH411132.1

Supplemental Table 3.2 Results of individual links parafit tests in viral family *Totiviridae*. For convience host phyla are abbreviated and host names are removed since viruses are only known from one host.

Host Phylum	Virus	F1.stat	p.F1	F2.stat	p.F2
Blasto	Allomyces.moniliformis.virus.4F	31.54836	0.009	1.38E-03	0.008
Blasto	Allomyces.sp-4C	88.83045	0.001	3.89E-03	0.001
Blasto	Allomyces.sp-4F	86.96015	0.001	3.81E-03	0.001
Asco	Alternaria.arborescens.victorivirus.1	73.15634	0.001	3.20E-03	0.001
Asco	Aspergillus.foetidus.slow.virus.1	57.92014	0.005	2.53E-03	0.004
Asco	Beauveria.bassiana.victorivirus.1	64.57654	0.003	2.83E-03	0.003
Asco	Botryosphaeria.dothidea.victorivirus.2	33.48614	0.012	1.47E-03	0.011
Asco	Colletotrichum.caudatum.totivirus.1	26.07816	0.012	1.14E-03	0.012
Asco	Conidiobolus.thromboides.virus.4A	28.58578	0.014	1.25E-03	0.013
Asco	Fusarium.poae.victorivirus.1	19.04568	0.025	8.33E-04	0.024
Asco	Gremmeniella.abietina.RNA.virus.L1	18.98883	0.023	8.31E-04	0.022
Asco	Helminthosporium.victoriae.virus.190S	-5.21657	0.897	-2.28E-04	0.898
Asco	Hortaea.werneckii.totivirus.1	-11.7305	0.985	-5.13E-04	0.987
Zoopag	Kickxella.alabastrina.virus.4D	-10.8944	0.99	-4.77E-04	0.99
Zoopag	Kickxella.alabastrina.virus.4G	12.649	0.056	5.54E-04	0.054
Zoopag	Kickxella.alabastrina.virus.4H	14.4564	0.065	6.33E-04	0.062
Asco	Nigrospora.oryzae.victorivirus.2	18.34432	0.032	8.03E-04	0.032
Chytrid	Operculomyces.laminatus.4A	28.67667	0.016	1.25E-03	0.016
Asco	Phomopsis.longicolla.totivirus.1	27.42494	0.016	1.20E-03	0.016
Asco	Phomopsis.vexans.RNA.virus	35.36065	0.015	1.55E-03	0.014
Basid	Puccinia.striiformis.totivirus.3	37.00821	0.006	1.62E-03	0.006
Basid	Puccinia.striiformis.totivirus.4	43.52359	0.009	1.90E-03	0.009
Asco	Rosellinia.necatrix.victorivirus.1	44.97389	0.011	1.97E-03	0.011
Asco	Saccharomyces.cerevisiae.virus.L-A	37.78868	0.01	1.65E-03	0.009
Asco	Saccharomyces.cerevisiae.virus.L-A-28	41.68861	0.011	1.82E-03	0.011
Asco	Saccharomyces.kudriavzevii.virus.L-A1	42.20089	0.006	1.85E-03	0.006
Asco	Saccharomyces.paradoxus.virus.L-A-21	44.72663	0.008	1.96E-03	0.008
Asco	Saccharomyces.paradoxus.virus.L-A-66	41.6703	0.009	1.82E-03	0.009
Asco	Scheffersomyces.segobiensis.virus.L	41.93769	0.015	1.84E-03	0.015

Asco	Sclerotinia.nivalis.victorivirus.1	19.14885	0.038	8.38E-04	0.036
Asco	Sphaeropsis.sapinea.RNA.virus.2	18.74547	0.028	8.20E-04	0.027
Asco	Thelebolus.microsporus.totivirus.1	23.28532	0.023	1.02E-03	0.023
Asco	Tolypocladium.ophioglossoides.totivirus.1	25.13446	0.028	1.10E-03	0.027
Asco	Tuber.aestivum.virus.1	25.45116	0.034	1.11E-03	0.032
Mucoro	Umbelopsis.nana.virus.4A	-11.0821	0.968	-4.85E-04	0.971
Mucoro	Umbelopsis.ramanniana.virus.2	1.722784	0.418	7.54E-05	0.414
Mucoro	Umbelopsis.ramanniana.virus.4	1.648424	0.442	7.21E-05	0.438
Asco	Ustilaginoidea.virens.RNA.virus.5	3.242251	0.364	1.42E-04	0.359
Basid	Xanthophyllomyces.dendrorhous.virus.L1A	2.356106	0.346	1.03E-04	0.341
Basid	Xanthophyllomyces.dendrorhous.virus.L1B	1.898789	0.391	8.31E-05	0.388
Chytrid	Zopfochytrium.polystomum.virus.4C	1.213781	0.465	5.31E-05	0.464
Chytrid	Cladochytrium.sp..JEL861.virus.4A	32.29806	0.02	1.41E-03	0.02
Chytrid	Cladochytrium.sp..JEL861.virus.4E	32.79743	0.017	1.44E-03	0.017

Supplemental Table 3.3 Results of individual links parafit tests in viral family *Partitiviridae*.

Host Phylum	Virus	F1.stat	p.F1	F2.stat	p.F2
Asco	Alternaria.alternata.partitivirus.1	12.12073	3.80E-02	8.17E-04	0.037
Asco	Aspergillus.flavus.partitivirus.1	12.82569	3.30E-02	8.65E-04	0.032
Asco	Aspergillus.fumigatus.partitivirus.1	15.47753	3.70E-02	1.04E-03	0.037
Asco	Beauveria.bassiana.partitivirus.1	12.5172	5.00E-02	8.44E-04	0.048
Asco	Botryosphaeria.dothidea.virus.1	7.863199	2.03E-01	5.30E-04	0.195
Mucoro	Choanephora.cucurbitarum.virus-2A	5.707809	2.78E-01	3.85E-04	0.274
Chytrid	Cladochytrium.JEL861.virus-2A	31.78828	3.90E-02	2.14E-03	0.039
Asco	Colletotrichum.acutatum.RNA.virus.1	90.42952	1.00E-03	6.10E-03	0.001
Asco	Colletotrichum.gloeosporioides.partitivirus.1	61.73949	5.00E-03	4.16E-03	0.005
Asco	Colletotrichum.truncatum.partitivirus.1	8.158688	0.099	5.50E-04	9.90E-02
Asco	Cytospora.sacchari.partitivirus.MP-2014	9.482865	0.075	6.39E-04	7.20E-02
Asco	Discula.destructiva.virus.1	-0.92557	0.916	-6.24E-05	9.17E-01
Asco	Discula.destructiva.virus.2	-1.02548	0.926	-6.91E-05	9.27E-01
Asco	Fusarium.solani.virus.1	-1.50235	0.929	-1.01E-04	9.29E-01
Asco	Gremmeniella.abietina.RNA.virus.MS1	6.676712	0.021	4.50E-04	2.10E-02
Asco	Gremmeniella.abietina.RNA.virus.MS1	8.178334	0.013	5.51E-04	1.30E-02
Asco	Gremmeniella.abietina.RNA.virus.MS2	11.88958	0.024	8.02E-04	2.40E-02
Asco	Magnaporthe.grisea.partitivirus.1	14.27489	0.019	9.62E-04	1.90E-02
Asco	Magnaporthe.oryzae.partitivirus.1	13.25901	0.013	8.94E-04	1.30E-02
Asco	Magnaporthe.oryzae.partitivirus.2	11.6147	0.015	7.83E-04	1.50E-02
Asco	Magnaporthe.oryzae.partitivirus.3	10.11064	0.02	6.82E-04	2.00E-02
Neocall	Neocallimastix.californiae.virus-2A	9.672993	0.015	6.52E-04	1.40E-02
Asco	Nigrospora.oryzae.partitivirus.1	13.57171	0.009	9.15E-04	9.00E-03
Asco	Ophiostoma.partitivirus.1	9.96907	0.022	6.72E-04	2.10E-02
Asco	Penicillium.aurantiogriseum.partitivirus.1	10.1406	0.021	6.84E-04	2.00E-02
Asco	Penicillium.brevicompactum.partitivirus.1	9.438726	0.022	6.36E-04	2.20E-02
Asco	Penicillium.digitatum.partitivirus.1	7.85448	0.026	5.30E-04	2.60E-02
Asco	Penicillium.stoloniferum.virus.F	9.242034	0.025	6.23E-04	2.50E-02
Asco	Penicillium.stoloniferum.virus.S	9.075221	0.018	6.12E-04	1.80E-02
Asco	Phyllosticta.citriasiana.partitivirus.1	7.080778	0.022	4.77E-04	2.20E-02

Asco	Pseudogymnoascus.destructans.partitivirus-pa	4.746595	0.033	3.20E-04	3.30E-02
Asco	Sclerotinia.sclerotiorum.partitivirus.3	4.081778	0.033	2.75E-04	3.20E-02
Asco	Sodiomyces.alkalinus.partitivirus.2	-0.64527	0.82	-4.35E-05	8.23E-01
Asco	Talaromyces.marneffei.partitivirus-1	-0.73226	0.828	-4.94E-05	8.32E-01
Asco	Thelebolus.microsporus.partitivirus.1	-0.90349	0.852	-6.09E-05	8.53E-01
Asco	Ustilaginoidea.virens.partitivirus	-2.52121	0.985	-1.70E-04	9.86E-01
Asco	Ustilaginoidea.virens.partitivirus	-2.44292	0.985	-1.65E-04	9.85E-01
Asco	Ustilaginoidea.virens.partitivirus.2	10.8912	0.042	7.34E-04	4.20E-02
Asco	Ustilaginoidea.virens.partitivirus.3	10.90057	0.059	7.35E-04	5.70E-02
Asco	Ustilaginoidea.virens.partitivirus.S2	10.80211	0.036	7.28E-04	3.40E-02

Supplemental Table 3.4 Results of individual links parafit tests in viral family *Mitoviridae*.

Host Phylum	Virus	F1.stat	p.F1	F2.stat	p.F2
Basid	Agaricus.bisporus.mitovirus.1	45.74609	1.00E-03	1.35E-03	0.001
Asco	Alternaria.arborescens.mitovirus.1	34.00012	3.00E-03	1.00E-03	0.003
Asco	Aspergillus.fumigatus.mitovirus.1	42.75631	2.00E-03	1.26E-03	0.002
Asco	Botrytis.cinerea.mitovirus.1	49.35378	2.00E-03	1.46E-03	0.002
Asco	Botrytis.cinerea.mitovirus.1.S	45.54688	3.00E-03	1.35E-03	0.003
Asco	Botrytis.cinerea.mitovirus.3	42.04845	2.00E-03	1.24E-03	0.002
Asco	Botrytis.cinerea.mitovirus.4	37.77137	2.00E-03	1.12E-03	0.002
Asco	Buergenerula.spartinae.mitovirus.1	39.0097	2.00E-03	1.15E-03	0.002
Basid	Ceratobasidium.mitovirus.A	37.39062	1.00E-03	1.10E-03	0.001
Basid	gi1024325058refYP_009249807.1	37.26034	0.003	1.10E-03	0.003
Asco	Cryphonectria.cubensis.mitovirus.2b	26.40051	0.005	7.80E-04	0.004
Asco	Sclerotinia.homoeocarpa.mitovirus	29.1671	0.001	8.62E-04	0.001
Asco	Colletotrichum.fructicola.mitovirus.1	12.66296	0.028	3.74E-04	0.028
Asco	Cryphonectria.parasitica.mitovirus.1-NB631	-2.99323	0.97	-8.85E-05	0.971
Asco	Leucostoma.persoonii.mitovirus.1	-3.45866	0.972	-1.02E-04	0.973
Asco	Leucostoma.persoonii.mitovirus.2	-3.19551	0.967	-9.44E-05	0.97
Zoopag	*Entomophthora.muscae.mitovirus.1*	-5.14142	0.999	-1.52E-04	0.999
Zoopag	*Entomophthora.muscae.mitovirus.2*	-3.81991	0.999	-1.13E-04	0.999
Zoopag	*Entomophthora.muscae.mitovirus.3*	22.48808	0.008	6.65E-04	0.007
Zoopag	*Entomophthora.muscae.mitovirus.4*	24.40096	0.011	7.21E-04	0.011
Zoopag	*Entomophthora.muscae.mitovirus.5*	22.76714	0.009	6.73E-04	0.008
Zoopag	*Entomophthora.muscae.mitovirus.6*	23.10255	0.006	6.83E-04	0.006
Zoopag	*Entomophthora.muscae.mitovirus.7*	23.10255	0.006	6.83E-04	0.005
Zoopag	*Entomophthora.muscae.mitovirus.8*	23.31901	0.006	6.89E-04	0.006
Asco	Erysiphe.necator.mitovirus.1	21.75221	0.01	6.43E-04	0.01
Asco	Erysiphe.necator.mitovirus.2	25.34118	0.017	7.49E-04	0.016
Asco	Erysiphe.necator.mitovirus.3	26.10259	0.027	7.71E-04	0.025
Asco	Fusarium.boothii.mitovirus.1	16.77526	0.019	4.96E-04	0.018
Asco	Fusarium.circinatum.mitovirus.1	15.21322	0.009	4.50E-04	0.008
Asco	Fusarium.circinatum.mitovirus.2-1	17.23404	0.001	5.09E-04	0.001

Asco	Fusarium.circinatum.mitovirus.2-2	17.23404	0.008	5.09E-04	0.008
Asco	Fusarium.coeruleum.mitovirus.1	25.35243	0.006	7.49E-04	0.005
Asco	Fusarium.globosum.mitovirus.1	25.76853	0.004	7.61E-04	0.004
Asco	Fusarium.oxysporum.f..sp..dianthi.mitovirus.1	25.29491	0.006	7.47E-04	0.005
Asco	Geopora.sumneriana.mitovirus.1	25.54785	0.009	7.55E-04	0.009
Mucoro	Gigaspora.margarita.mitovirus.1	33.50956	0.005	9.90E-04	0.005
Mucoro	Gigaspora.margarita.mitovirus.2	27.83275	0.002	8.22E-04	0.002
Mucoro	Gigaspora.margarita.mitovirus.3	30.20504	0.003	8.93E-04	0.003
Mucoro	Gigaspora.margarita.mitovirus.4	28.86744	0.004	8.53E-04	0.004
Asco	Gremmeniella.mitovirus.S1	23.69007	0.006	7.00E-04	0.006
Asco	Helicobasidium.mompa.mitovirus.1-18	20.29779	0.002	6.00E-04	0.002
Basid	Heterobasidion.mitovirus.1	13.86358	0.014	4.10E-04	0.014
Asco	Hymenoscyphus.fraxineus.mitovirus.1	6.765701	0.119	2.00E-04	0.114
Asco	Leptosphaeria.biglobosa.mitovirus.1	7.424118	0.098	2.19E-04	0.093
Asco	Leucostoma.cinctum.mitovirus.2	5.387613	0.157	1.59E-04	0.154
Asco	Macrophomina.phaseolina.mitovirus.2	5.706176	0.171	1.69E-04	0.165
Asco	Macrophomina.phaseolina.mitovirus.3	5.702115	0.136	1.69E-04	0.132
Mucoro	Mortierella.minutissima.virus-1C	5.635482	0.215	1.67E-04	0.212
Mucoro	Mortierella.multidivaricata.virus-1A	6.243446	0.146	1.84E-04	0.142
Mucoro	Mortierella.verticillata.virus-1A	6.953321	0.17	2.05E-04	0.168
Mucoro	Mortierella.verticillata.virus-1B	6.953321	0.179	2.05E-04	0.177
Asco	Ophiostoma.mitovirus.3a	11.28829	0.045	3.34E-04	0.044
Asco	Ophiostoma.mitovirus.4	10.57224	0.029	3.12E-04	0.026
Asco	Ophiostoma.mitovirus.5	11.45301	0.042	3.38E-04	0.04
Asco	Ophiostoma.mitovirus.6	9.316814	0.032	2.75E-04	0.031
Asco	Ophiostoma.mitovirus.7	8.969853	0.049	2.65E-04	0.047
Basid	Puccinia.striiformis.mitovirus.1	7.803632	0.059	2.31E-04	0.058
Basid	Rhizoctonia.mitovirus.1	4.683577	0.154	1.38E-04	0.154
Basid	Rhizoctonia.mitovirus.1.RS006-2	4.697122	0.172	1.39E-04	0.169
Basid	Rhizoctonia.mitovirus.2.RS002	20.83102	0.467	6.16E-04	0.457
Mucoro	Rhizophagus.irregularis.mitovirus.1	0.260844	0.839	7.71E-06	0.839
Mucoro	Rhizophagus.irregularis.mitovirus.2	1.909804	0.449	5.64E-05	0.441
Asco	Sclerotinia.sclerotiorum.mitovirus.1.HC025	1.782266	0.573	5.27E-05	0.56

Asco	Sclerotinia.sclerotiorum.mitovirus.12-A	1.039541	0.626	3.07E-05	0.62
Asco	Sclerotinia.sclerotiorum.mitovirus.14-A	9.774322	0.068	2.89E-04	0.063
Asco	Sclerotinia.sclerotiorum.mitovirus.2	9.330297	0.063	2.76E-04	0.058
Asco	Sclerotinia.sclerotiorum.mitovirus.27	8.867621	0.039	2.62E-04	0.038
Asco	Sclerotinia.sclerotiorum.mitovirus.28	8.929967	0.047	2.64E-04	0.046
Asco	Sclerotinia.sclerotiorum.mitovirus.29	10.67276	0.049	3.15E-04	0.047
Asco	Sclerotinia.sclerotiorum.mitovirus.30	11.29463	0.048	3.34E-04	0.046
Asco	Sclerotinia.sclerotiorum.mitovirus.31	11.20082	0.042	3.31E-04	0.041
Asco	Sclerotinia.sclerotiorum.mitovirus.32	11.01753	0.042	3.26E-04	0.04
Asco	Sclerotinia.sclerotiorum.mitovirus.33	22.01768	0.101	6.51E-04	0.097
Asco	Sclerotinia.sclerotiorum.mitovirus.34-A	15.00237	0.043	4.43E-04	0.043
Asco	Sclerotinia.sclerotiorum.mitovirus.4	13.59985	0.052	4.02E-04	0.051
Asco	Sclerotinia.sclerotiorum.mitovirus.5-A	12.69856	0.036	3.75E-04	0.035
Asco	Sclerotinia.sclerotiorum.mitovirus.6	11.45077	0.032	3.38E-04	0.031
Asco	Sclerotinia.sclerotiorum.mitovirus.7-A2	11.80056	0.041	3.49E-04	0.038
Asco	Sclerotinia.sclerotiorum.mitovirus.8-A	5.529055	0.208	1.63E-04	0.201
Zoopag	Syncephalis.fuscata.virus-1A	4.328378	0.257	1.28E-04	0.251
Zoopag	Syncephalis.fuscata.virus-1B	4.306236	0.246	1.27E-04	0.243
Asco	Thielaviopsis.basicola.mitovirus	26.63302	0.001	7.87E-04	0.001
Asco	Tuber.aestivum.mitovirus	30.37831	0.002	8.98E-04	0.002
Asco	Tuber.excavatum.mitovirus	28.87112	0.005	8.53E-04	0.005
Asco	Valsa.cypri.mitovirus	45.24233	0.002	1.34E-03	0.002

Supplemental Table 3.5 Results of individual links parafit tests in viral family *Chrysoviridae*.

Host Phylum	Virus	F1.stat	p.F1	F2.stat	p.F2
Asco	Alternaria.alternata.chrysovirus.1	51.15835	0.01	8.23E-03	0.01
Asco	Aspergillus.coremiiformis.virus	7.331745	0.182	1.18E-03	0.182
Asco	Aspergillus.fumigatus.chrysovirus	0.477384	0.457	7.68E-05	0.453
Asco	Aspergillus.mycovirus.1816	0.432664	0.451	6.96E-05	0.448
Asco	Aspergillus.thermomutatus.chrysovirus.1	-0.09274	0.646	-1.49E-05	0.647
Asco	Beauveria.bassiana.chrysovirus.1	6.865277	0.033	1.10E-03	0.032
Asco	Bipolaris.maydis.chrysovirus.1	6.177348	0.039	9.94E-04	0.038
Asco	Botryosphaeria.dothidea.chrysovirus	10.58273	0.026	1.70E-03	0.026
Asco	Chrysothrix.chrysovirus.1	11.6182	0.018	1.87E-03	0.018
Asco	Colletotrichum.gloeosporioides.chrysovirus.1	8.862955	0.014	1.43E-03	0.013
Asco	Coniothyrium.diplodiella.chrysovirus.1	9.478289	0.02	1.52E-03	0.019
Asco	Cryphonectria.nitschkei.chrysovirus.1	8.90695	0.023	1.43E-03	0.022
Asco	Fusarium.graminearum.dsRNA.mycovirus.2	15.37346	0.019	2.47E-03	0.018
Asco	Fusarium.graminearum.mycovirus.China.9	15.53228	0.014	2.50E-03	0.014
Asco	Fusarium.oxysporum.chrysovirus.1	7.63056	0.04	1.23E-03	0.04
Asco	Helminthosporium.victoriae.145S.virus	5.144235	0.087	8.27E-04	0.085
Asco	Isaria.javanica.chrysovirus.1	4.045883	0.061	6.51E-04	0.061
Asco	Macrophomina.phaseolina.chrysovirus.1	4.612585	0.061	7.42E-04	0.06
Asco	Neofusicoccum.parvum.chrysovirus.1	4.249818	0.061	6.84E-04	0.06
Asco	Penicillium.chrysogenum.virus	4.295353	0.056	6.91E-04	0.055
Asco	Penicillium.italicum.chrysovirus	3.947156	0.059	6.35E-04	0.059
Asco	Penicillium.janczewskii.chrysovirus.1	4.58906	0.06	7.38E-04	0.058
Asco	Penicillium.janczewskii.chrysovirus.2	4.394279	0.061	7.07E-04	0.06
Asco	Penicillium.raistrickii.chrysovirus.1	3.99409	0.065	6.42E-04	0.065
Asco	Penicillium.roseopurpureum.chrysovirus.1	3.947675	0.074	6.35E-04	0.073
Asco	Tolypocladium.cylindrosporum.virus.2	0.738759	0.342	1.19E-04	0.337
Asco	Verticillium.dahliae.chrysovirus.1	0.785663	0.275	1.26E-04	0.272
Basid	White.button.mushroom.virus.1	0.791709	0.278	1.27E-04	0.272
Mucoro	Mortierella.minutissima.virus-4A	2.204323	0.094	3.55E-04	0.092
Blasto	Allomyces.sp.virus-4E	2.170421	0.104	3.49E-04	0.103

Blasto	Allomyces.moniliformis.virus-4A	1.538153	0.14	2.47E-04	0.139
Blasto	Allomyces.sp.virus-4B	3.696924	0.343	5.95E-04	0.341

Chapter 4 Basal fungi are host to large DNA viruses

Abstract

As their name suggests, the nucleocytoplasmic large DNA viruses (NCLDVs) are large viruses with unique replication strategies that are thought to have origins coinciding with or preceding the diversification of eukaryotes. Consistent with this hypothesis, they are hosted by a diverse suite of organisms, from insects to algae to protists. A kingdom for which there is no known host of NCLDVs, however, is Fungi. Recently, genomic and metagenomic evidence suggested the occurrence of NCLDVs of fungi, but these studies were unable to assign hosts nor distinguish viral genes that had become endogenous in the host genomes versus independently replicating viruses. We sequenced, assembled, and searched the genomes of 93 early-diverging fungi for evidence of infection by NCLDVs. We found 16 genomes containing the "smoking gun" of NCLDVs, the major capsid protein, and at least 7 isolates with most of the core NCLDV genes. These viruses form a well-supported clade related to *Mimiviridae* and *Phycodnaviridae*; we propose a new family *Mycodnaviridae.* We speculate that these viruses are the same "gamma particles" from historic records of members of the Blastocladiomycota. Like the *Phycodnaviridae*, the replication strategy of *Mycodnaviridae* may involve active replication in the motile, cell wall-less gametes and subsequent endogenization in the host genome. Fungal NCLDVs, therefore, likely only occur in fungi with life histories that include a zoosporic stage. Thus, it is unsurprising that fungal NCLDVs have not been previously discovered, since these early-diverging fungal lineages are grossly understudied. Overall, our study has expanded the

known breadth of NCLDV host organisms and yielded exciting new model systems with which to further probe viral-host interactions.

Introduction

The most abundant and diverse DNA viruses of eukaryotes are the nucleo-cytoplasmic large DNA viruses (NCLDVs) (Koonin and Yutin 2010). This group was pragmatically dubbed this common name for the location of transcription, which is either the cytoplasm, the nucleus, or beginning in the nucleus and completing in the cytoplasm, and for their large genomes (Iyer 2001). The original members were *Poxviridae, Phycodnaviridae, Asfarviridae,* and *Iridoviridae*, which form a monophyletic group composed of double-stranded DNA linear genomes ranging from ~100-500kb (Iyer 2001).

These viral groups display a range of host interactions and consequential societal implications. African swine fever virus (*Asfarviridae)* is a major agricultural concern as it causes deadly hemorrhagic fevers in domesticated pigs and yet is apparently asymptomatic and persistent in its natural suid reservoir hosts (reviewed in Galindo and Alonso 2017). *Poxviridae* hosts include humans, other vertebrates, and arthropods. One celebrated poxvirus, Vaccinia virus, is avirulent in humans and provides immunity to another, deadly, poxvirus, Variola virus, the causative agent of smallpox. Thanks to vaccinations with live Vaccinia virus, smallpox has been eradicated in nature. *Phycodnaviridae* are a speciose group of viruses that infect algae and have particularly multifarious replication strategies: chloroviruses are lytic and encode numerous cell wall-degrading enzymes while the phaeoviruses are lysogenic and only infect the free-swimming and cell wall-less gametes and zoospores of their hosts (reviewed in Wilson 2009). *Coccolithovirus* infection of the calcified photosynthetic protist *Emiliana huxleyi* during blooms

has recently been recognized as a major contributor to the remineralization of the mesopelagic zone of the North Atlantic (Laber 2018), demonstrating the significant contribution of this NCLDV to global nutrient cycling.

All viruses must confront the conundrum of diverting host resources for viral purposes, which creates conflict with the host and competition for transcriptional machinery. NCLDVs have a unique approach to circumvent this issue: they provide their own transcriptional apparatus and reorganize the host cytoplasm to create a compartmentalized "viral factory" for virion production. These viral factories typically feature specialized locales for viral transcription and protein synthesis, effectively concentrating viral proteins for efficiency (Condit 2007). Interestingly, these viral factories are proposed to have played a central role in eukaryogenesis and the origin of the nucleus (Forterre and Gaïa. 2016), which coincides with hypotheses that NCLDVs have ancient origins predating or coinciding with the diversification of eukaryotes (Guglielmini 2019).

Because of the unique replication strategies, societal and ecological relevance, many of these NCLDVs had been studied for decades; but in 2003, the field of virology was turned upside-down by the discovery of a new NCLDV— *Acanthamoeba polyphaga* Mimivirus (APMV) (La Scola 2003). Far greater in both particle and genome size than any known virus at the time(Raoult 2004), and larger than some single-celled bona fide organisms, this finding effectively reinvigorated the debate of a possible fourth domain of life from which this "giant virus" reduced (Boyer 2010, Colson 2011, Legendre 2012, Yutin 2014, Nasir 2012, Williams 2011, Moreira 2015). Following the discovery of APMV was a scientific surge to isolate and characterize related giant viruses. The push to find more–and bigger–viruses led to the isolation of hundreds of NCLDVs, particularly by an amoeba reporter assay (Boughalmi 2013), and has

led to significant gains in our understanding of the diversity, ubiquity, and structure of NCLDVs. However, the reliance on a reporter assay contributed to a lack of discovery in other hosts, and thus the underlying biology of many NCLDVs remains shrouded by a lack of observation within the natural host.

The currently known NCLDVs are highly diverse genetically and morphologically. They range in genome size from ~100 kb (the smallest iridoviruses) to the 2.5 MB pandoraviruses. Structure varies from icosahedral to ovoid to spherical to globular. While the official taxonomy according to the International Committee on the Taxonomy of Viruses (ICTV) is currently in flux, the most recent framework was built with the recognition that the identified viruses in these groups are expanding at near-exponential rates (Koonin 2019). These currently include seven families (*Mimiviridae, Phycodnaviridae, Ascoviridae, Iridoviridae, Poxviridae, Asfarviridae, and Marseilleviridae)* and have been tentatively classified into 5 orders and two classes within the phylum *Nucleocytoviricota*, kingdom *Bamfordvirae*, realm *Varidnaviria* (Koonin 2019).

Since the discovery of APMV, the "fourth domain" debate has been quelled, as evidence supports the origin of NCLDVs as a prokaryotic DNA virus, with subsequent accumulation of genes from a variety of hosts and substantial gene duplication in some lineages (Yutin 2014, Filée 2013, Koonin 2015, Suhre 2005). Gene accumulation by NCLDVs has yielded dynamic genomes that form complex phylogenetic networks (Moniruzzaman 2020), but nonetheless abundant evidence supports their common ancestry (Iyer 2001, Yutin 2014, Koonin 2019, Iyer 2006). Since the first giant virus discovery, there have been many NCLDV phylogenomic studies, largely driven by metagenomic sequencing analyses (Moniruzzaman 2020, Hingamp 2013, Moniruzzaman 2017, Schulz 2020).

These metagenomic analyses have revealed the surprising abundance and diversity of globally- distributed NCLDVs. Leveraging the plenitude of quality marine metagenomic datasets, studies have shown up to 10^5 NCLDV genomes can exist within a single microliter of ocean water (Hingamp 2013) and the taxon richness of just one NCLDV order, *Imitervirales* (formerly proposed *Megavirales),* surpasses that of both bacteria and archaea in oceans (Mihara 2018). Among the environments tested, datasets from aquatic environments have consistently been found to have the greatest numbers of NCLDV genomes (Moniruzzaman 2020, Schulz 2020), perpetuating the idea that NCLDVs are most often associated with aquatic organisms. Metagenomic datasets from terrestrial environments have yielded markedly fewer NCLDVs (Moniruzzaman 2020, Schulz 2020) but enrichment techniques such as "mini-metagenomics" has revealed previously hidden diversity in soils and hints that much diversity in these habitats remains to be uncovered (Schulz 2018). Metagenomics has proven a powerful tool for NCLDV discovery, but a major limitation is the inability to directly associate viruses with their hosts; though, recently developed inference techniques are promising to overcome this (Moniruzzaman 2017, Schulz 2020).

Confirmed NCLDV-host associations include hosts from a breadth of taxa— from single-celled algae to invertebrates to animals (reviewed in Koonin 2015)— and metagenomic- and metatranscriptomic-powered inferences suggest an even broader scope of eukaryotic hosts (Moniruzzaman 2017, Schulz 2020). Interestingly, there is evidence of horizontal gene transfer of fungal genes into numerous NCLDV lineages (Schulz 2020, Schulz 2018), though no fungi have been previously reported to be host to NCLDVs. Much of what is known about NCLDV biology has been learned from protist-NCLDV interactions. The development of additional laboratory model systems of NCDLVs and their natural hosts could lead to advances in our

understanding of virus-host interactions, including host immune response, viral "reprogramming" of host cells, metabolic impacts, and their ecosystem-level consequences.

We describe here the NCLDVs found infecting fungi, a kingdom previously unknown to host NCLDVs. We searched 93 genomes of early-diverging fungi for evidence of NCLDVs and uncovered a lineage of viruses in the fungal phyla Blastocladiomycota and Chytridiomycota, for which we propose a new viral family *Mycodnaviridae.* These viruses have an average genome size of approximately 300kb, appear to have a core suite of at least 40 genes, and are related to *Mimiviridae* and *Phycodnaviridae*. We speculate that historically reported self-replicating and transmissible organelles present in some members of the Blastocladiomycota, termed "gamma particles", are in fact large viruses. If true, it is likely that *Mycodnaviridae* replicate via a lysogenic strategy and, perhaps, occur exclusively in the basal fungal lineages. While infected isolates are apparently asymptomatic, we predict there are metabolic implications to infection consistent with other NCLDVs. Our (re)identification of an historically reported curiosity is a humbling reminder of the unsung scientific feats of the past, and an invigorating call to future study.

Methods

Fungal cultivation, DNA extraction, and whole genome sequencing

The genomic data for this study were primarily generated by the Joint Genome Institute, with an additional 12 strains sequenced at the University of Michigan (Supplementary Table 4.1) and supplemented with 5 genomes obtained from GenBank (van de Vossenberg 2019). Almost all of the JGI genomes were generated from nucleic acids provided by our research group. JGI genomes were sequenced at either high depth of coverage (>100X, n=32) by either Illumina or

PacBio single molecule sequencing or at low depth of coverage (n=49) by Illumina. Details of sequencing and assembly of JGI genomes have been published (Chang 2015, Ahrendt 2018, Mondo 2017). Most cultures are available from the Collection of Zoosporic Eufungi at the University of Michigan (CZEUM) (Simmons 2020). We grew isolates in media appropriate for their nutritional needs (see protocols at https://czeum.herb.lsa.umich.edu), harvested tissue, and extracted DNA using the CTAB method (James 2008). For genomes sequenced at the University of Michigan, we prepped libraries for whole genome sequencing by using the Nextera XT kit to produce paired end 150 bp reads on an Illumina HiSeq 4000.

Genome assembly and annotation

Assemblies were conducted using SPAdes 3.11.1 (Nurk 2013). Because many of the nuclear genomes were believed to be diploid, we assembled the genomes using both dipSPAdes (Safonova 2014), which assumes a diploid genome and tries to merge sequences from two alleles into a single contig, as well as the normal haploid mode of SPAdes. Assemblies that were considerably smaller >10% difference in size in dipSPAdes vs. SPAdes and had higher genome continuity (e.g., N50) were assumed to be diploid and the dipSPAdes assemblies used. In order to preserve small fragments that may be viral we did not filter assemblies and only removed contigs that were less than <500 bp in size. Genomes were separately annotated as fungi using funannotate (DOI: 10.5281/zenodo.2604804) and as virus using Prodigal (Hyatt 2010). We searched the annotated genomes with HMMER2 (Eddy 2009) using consensus sequences made from NCVOGs in Yutin 2014. Using the Prodigal annotations, we computed several statistics on each of the contigs in an attempt to distinguish the minority viral contigs from fungal. These included %GC, % of ORFs matching NCLDV orthologs, gene density and intergenic distance,

and coverage. We retrieved annotated genomes of reference NCLDVs from RefSeq (Supplementary Table 4.2).

Identifying viral homologs

We searched fungal and reference annotated genomes with HMMER2 (Eddy 2009) as in Schulz 2020: using hmms of a subset of the nucleocytoplasmic viral orthologous genes (NCVOGs)— the 20 NCVOGs most likely to have been vertically inherited (Yutin 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 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 2009) with 100 bootstrap replicates, rooting at the midpoint.

We used these gene trees in combination with the program ViralRecall (https://github.com/faylward/viralrecall) to discern putative viral genes from fungal homologs. ViralRecall uses HMMs to detect viral signatures in genomic data relative to cellular genomic signals. The two databases used by ViralRecall are Pfam (modified by removing common viral protein families) and VOGDB, a comprehensive database of all viral proteins in RefSeq. Both stringent (-s 15 -m 30 -v 10) and relaxed (-s 15 -m 2 -v 2) parameters were used, with the latter used for poorer quality genomes assemblies. We considered an NCVOG hit from a fungal genome as valid if:

- found on same contig as a major capsid protein, or
- having viral regions making up $\geq$ 70% of the total contig with a ViralRecall score $\geq$ 25 AND
    - encoding a protein forming a clade with *only* reference NCLDV genes or fungi with MCPs (ex. Supplementary Figures 4.4 and 4.7), or

- encoding a protein forming a clade with *primarily* genes of reference NCLDVs and/or fungi with MCPs, where other genes in that clade are found on the same contig as the MCP (ex. Supplementary Figure 4.10)

One isolate included the signature of what appeared to be two distinct NCLDVs; we used MetaBAT2 (Kang 2019) on the entire genome of *Allomyces javanicus* California12 under the default settings to determine whether coinfection was occurring and to separate the two NCLDV genomes.

Concatenated NCLDV phylogenetic trees

We used 5 conserved NCVOGs in the phylogenetic reconstruction of the NCLDVs: NCVOG0022 (Major Capsid Protein), NCVOG0023 (D5-like helicase primase), NCVOG0038 (DNA polymerase elongation subunit family B), NCVOG0076 (DNA or RNA helicases of superfamily II), and NCVOG0249 (A32-like packaging ATPase). Where genomes had multiple copies of an NCVOG, we chose the hit with the lowest e-value. We used ProtTest (Darriba 2011) to determine the best-fit model of amino acid substitution for the concatenated alignments and reconstructed phylogenetic trees with the maximum-likelihood approach implemented in RaxML (Stamatakis 2014) (-f a) with 100 bootstrap replicates.

Genome comparisons

We considered the viral genomes of *Allomyces arbuscula* Burma1F, *Allomyces javanicus* California12, and *Blyttiomyces helicus* to consist of all contigs in which the viral region made up $\geq$ 70% of the total and had a score $\geq$25 per ViralRecall, with the exception of *Allomyces javanicus* California12. Because this isolate is co-infected with two NCLDVs, we used the results of metaBAT2 and selected one of the two viral bins to include in this analysis; the single contig making up the selected genome was validated by ViralRecall. We conducted TBLASTX searches of each viral genome against the others and filtered the results by alignment length

(≥50bp), bitscore (≥40), and e-value (<0.001). We calculated the GC content of contigs with a sliding window of 500bp. We conducted BLASTP searches of each prodigal-predicted proteome against nr to identify homologs (e-value <1e-10). Because many symbiont genes are likely mis-identified 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 hits were only to the fungi which we have shown encode MCP; though these are likely viral genes, we conservatively labeled these as "strictly fungal" since relatively few whole genomes of early-diverging fungi exist in the databases currently. We visualized these results using the program Circos (Krzywinski 2009).

We used OrthoVenn2 (Xu 2019) to compare orthologous gene clusters of these three viral genomes' Prodigal annotations (e-value <1e-5; inflation value 1.5). Protein sequences of the gene clusters shared by the three viruses were functionally annotated using eggNOG-mapperv2 (Huerta-Cepas 2017, Huerta-Cepas 2019) under default parameters (max e-value 0.001; min bitscore 60, min % of query cover 20) as were the entire protein coding regions of each virus.

Results

We identified homologs of NCVOG0022, the NCLDV major capsid protein (MCP), in 16 fungal (meta)genomes (Table 4.1). Of those, 9 genomes have at least 3 of the 5 core NCVOGs (NCVOG0022 (MCP), NCVOG0023 (D5-like helicase primase), NCVOG0038 (DNA polymerase elongation subunit family B), NCVOG0076 (DNA or RNA helicases of superfamily II (COG1061)), NCVOG0249 (A32-like packaging ATPase)) and 7 genomes have at least 4 of the 5 core NCVOGs (Table 4.1; Figure 4.1; Supplementary Figures 4.1-4.12). For 3 of the 16 isolates (*Fimicolochytrium jonesii* JEL569, *Gaertneriomyces semiglobifer* Barr 43, and *Chytriomyces hyalinus* ARG121), the only core NCVOG definitively identified was the major capsid protein. These MCP homologs are on contigs which do not pass the viral cutoffs per

ViralRecall, so it is possible that these species historically had infections, and the MCPs genes have been transferred to the fungal genome. Indeed, we identified two additional isolates of *Chytriomyces hyalinus* with evidence of NCLDV infection. Overall, the number of isolates with most of the core NCVOGs is likely an underestimate, as genome assemblies varied in quality.

The 16 genomes with evidence of present or historic NCLDVs are in Chytridiomycota (n=13) and Blastocladiomycota (n=3). Although the Blastocladiomycota are particularly poorly sampled, and thus our genomic searches were limited (n=6), NCLDVs were found only in the Blastocladiomycetes. Among the infected members of Chytridiomycota, the Chytridiaceae (Chytridiomycetes, Chytridiales) were highly represented (n=6), followed by the Gonapodaceae (Monoblepharidomycetes, Monoblepharidales; n=2) and the Spizellomycetales (Spizellomycetes; n=2). The Cladochytriales and Rhizophlyctidales each had one infected genome, and one isolate is yet unplaced at the order level. Thus, NCLDVs do not appear to be present in every class of these two phyla, and instead appear to have been lost in some lineages.

The seven fungal NCLDVs with at least 4 of the 5 core genes (hereafter referred to as *Allomyces javanicus DNA virus 1, Allomyces arbuscula DNA virus 1, Catenaria anguillulae DNA virus 1, Blyttiomyces helicus DNA virus 1, Chytriomyces hyalinus DNA virus 1, Quaeritorhiza hematococci DNA virus 1, Polychytrium aggregatum DNA virus 1)* form a well-supported clade (Figure 4.2). We propose the construction of a new family, *Mycodnaviridae* in class *Megaviricetes*, phylum *Nucleocytovaricota*, kingdom *Bamfordvirae*. The new group is sister to *Mimiviridae* according to this phylogeny, although without bootstrap support (37%, Figure 4.2). Some gene trees substantiate this result (MCP (NCVOG0022), Figure 4.1; A32-like packaging ATPase (NCVOG0249), Supplementary Figure 4.4; FtsJ-like methyltransferase

family protein (NCVOG0059), Supplementary Figure 4.7), though more in-depth characterization is needed to determine proper placement of *Mycodnaviridae*.

Comparison of three viral genomes reveals high levels of punctuated synteny, consistent with expectations of an ancient common ancestor (Figure 4.3). *Blyttiomyces helicus DNA virus 1*, *Allomyces arbuscula DNA virus 1*, and *Allomyces javanicus DNA virus 1* each have high percentages of unknown ORFs (39%, 57%, and 60%, respectively) and virus orthologs (19%, 18%, 20%, respectively). The *B. helicus* genome has been previously published and thus is in the database which skews these results; 20% of the *Blyttiomyces helicus DNA virus 1* genome hit only to fungi with MCPs, which largely comprises self-hits. Functionally, these three genomes are also largely composed of proteins with unknown function per eggNOG analysis (Figure 4.4), and, unsurprisingly, proteins for replication, recombination, and repair.

There appears to be a core suite of 43 gene clusters common to these three annotated viral genomes which could define the group. Overall, these genomes formed 177 gene clusters, 141 of which are orthologous clusters (at least two viral genomes represented) and 43 of which are shared by all three viruses (Figure 4.5). Of these 43 shared clusters, 27 have associated GO annotations (Supplementary Table 4.3). By functional characterization of the 43 shared clusters, we found 41% to have unknown function (Figure 4.6). Other clusters are involved in expected processes such as replication, recombination, and repair (13%), transcription (6%), translation, ribosomal structure, and biogenesis (1%), and post-translational modification, protein turnover, and chaperones (7%). The shared clusters also include proteins involved in cell cycle control, cell division, chromosome partitioning (2%), cell wall/membrane/envelope biogenesis (7%), intracellular trafficking, secretion, and vesicular transport (1%), and cytoskeleton (2%) which could be involved in formation of the NCLDV signature “viral factory” and virion assembly.

Especially interesting, 5% of the functional groups represented are involved in carbohydrate transport and metabolism, which, speculatively, could increase host fitness by increasing its metabolic flexibility.

The 7 viruses with at least 4/5 core genes have an average genome size of 308,478 bp, which includes all the contigs passing our ViralRecall criteria (Supplementary Table 4.3; the ViralRecall results for *Polychytrium aggregatum DNA virus 1* appear contaminated by fungal genes and was removed per Grubbs' test of outlier detection ($p < 0.05$)). Of the known NCLDVs, *Mycodnaviridae* are thus placed amongst the moderately large, and not giant. Estimates of genome sizes should be considered with some caution, however, as we have yet to sequence isolated particles.

One genome, *Allomyces javanicus* California-12, contains two distinct sets of NCVOGs and approximately double the viral genome size of that found in *Allomyces arbuscula*. Binning with Metabat2 confirmed the presence of two NCLDV bins, each with at least 3 of the 5 core NCVOGs. This instance of apparent natural coinfection is, to our knowledge, unprecedented. Interestingly, *Allomyces javanicus* California-12 appears to have arisen through hybridization (James, unpublished data). We speculate that the two parental strains of *Allomyces* were each host to unique NCLDVs that were maintained in the hybrid, which is plausible given the apparent consistency of viral infection in *Allomyces.*

Discussion

The story of large and giant virus identification is dynamic and ongoing. One common theme is the initial misidentification of large viruses as endosymbionts. In 1992, an obligate intracellular "bacterium" was identified in free-living amoebae isolated from the water of a cooling tower associated with a pneumonia outbreak. It took eleven years for the accurate

identification of this “endosymbiont” as *Acanthamoeba polyphaga Mimivirus* (La Scola 2003). A few years later, an “Archaea-like endosymbiont”, suspiciously similar to the later-described *Pithovirus sibericum*, was described in an amoeba (Hoffmann 1998). In 2008 Schied and colleagues discovered yet another amoebae “endosymbiont” that was later revealed to be one of the largest of the giant viruses to date, *Pandoravirus opinatum* (Scheid 2008, Scheid 2016). Given the exceptional size and morphological diversity of these viruses, it is not difficult to understand why initial misidentifications are common. The results of our study suggest yet another example of this theme.

Unusual “gamma particles” were first described in zoospores of the fungus *Blastocladiella emersonii* (phylum *Blastocladiomycota)* in 1956 (Cantino and Horenstein) and shown to possess DNA over a decade later (Cantino and Mack 1965, Myers and Cantino 1971). Similar self-replicating particles were demonstrated in additional isolates of *Blastocladiella* (Matsumae 1970) and the related genera *Allomyces* and *Catenaria* (Barstow 1979). The similarities to large viruses are striking. Barstow and Lovett (1975) describe the sequence of gamma particle formation in *B. emersonii*: 40nm electron dense granules appear in the cisternae of the rough endoplasmic reticulum (ER), then undergo stages of aggregation to finally form 300nm particles, though other studies describe sizes ranging up to 500nm (Cantino and Horenstein 1956). NCLDVs produce virions via association with the ER membrane; particularly, the internal virion envelope is formed through ER disrupting mechanisms (reviewed in Romero-Brey and Bartenschlager 2016). Cantino and Horenstein (1956) demonstrated the transferability, via cytoplasmic exchange, of the particles they would go on to designate as “gamma particles” and so thoroughly describe. Most peculiar at the time was the observation that gamma particles are present only in zoospores (Lessie and Lovett 1968). Upon zoospore encystment, the gamma

particles dissociate and seem to disappear in apparent concert with phases of encystment including retraction of the flagellum and dissolution of the nuclear cap (Truesdell and Cantino 1970). While the replicative cycle of these so-called "gamma particles" is unusual, it is not unprecedented.

Phaeoviruses *(Phycodnaviridae)* infect brown algae, which, like *Allomyces* and *Blastocladiella,* have complicated life cycles with alternation of generations: male and female haploid gametophytes produce swimming haploid zoospores, which fuse to form diploid sporophytes that undergo meiosis, producing haploid zoospores that give rise to the gametophytes. Phaeoviruses only infect the cell wall-less, swimming gametes or zoospores of their hosts, and are otherwise integrated in the host genome where they are preserved through mitotic division. When the filamentous host develops sexual structures, viral replication is triggered, and virions are released to infect free-swimming gametes and begin the cycle again (Wilson 2009).

Though confirmatory work is needed, it is tantalizing to consider that these gamma particles, subject to thorough scientific inquiry in the 1960s and 1970s and yet still so mysterious, could in fact be the large viruses we have identified in genomic data. The similarity in replication to the phaeoviruses is undeniable, perhaps suggesting a common strategy of infection in organisms whose life cycles include a cell wall-less zoospore stage. The finding of two seemingly independent NCLDVs in one isolate, the predicted hybrid *A. javanicus* California-12 is particularly compelling given this hypothesis of lysogeny, if conjectural. Future work will test whether integration of the virus into the host genome is occurring. Putting to rest the mystery of the "gamma particle" is a potent reminder of the wealth of knowledge that remains dormant in the past. It simultaneously incites new questions into the functional

ramifications of NCLDVs to fungal evolution and to the ecology of these inconspicuous organisms.

Conclusion

It is hardly surprising to find evidence of large viruses in fungi, since NCLDVs are geographically and taxonomically widespread, wildly diverse, and their study is only just coming of age. However, the study of DNA viruses in fungi generally is surprisingly limited. In fact, only two DNA viruses have been identified in the entire kingdom to date (Li 2020, Yu 2010), and these are both small (2-6 kb), circular rep-encoding single-stranded (CRESS) DNA viruses. And so, while it may not be surprising to uncover NCLDV evidence in fungi, it is an exciting contribution to the understanding of DNA virus infection in fungi generally.

Additionally, it is unsurprising that fungal NCLDVs had not been identified (as viruses) until now, since infection is likely restricted to the basal fungal lineages, which have been poorly studied. Further, these viruses do not appear to have a recognizable phenotype and are persistent through years of culturing. Attempts to cure infected strains have yet to yield virus-free isolates. It is all the more interesting that these basal lineages are where NCLDVs are found since their life cycles often depend on the presence of water. Many of these organisms are entirely aquatic and others are soil-dwellers, but all have a motile, zoosporic stage. It has become clear from metagenomic studies that NCLDVs are most abundant in aquatic environments. Our findings are consistent with this, although the biological meaning remains mysterious.

Recent work demonstrating the “physiological reprogramming” of host cells in response to NCLDV infection (Moniruzzaman 2018) has made it clear that metabolomics holds a key for disentangling the role of NCLDV infection in large-scale ecological processes. Nanoscale-drivers of ecological processes have been recognized as penultimate in marine ecosystems

(Brussaard 2008). Our work contributes to a body of literature forming that invariably shows the same to also be true in freshwater and terrestrial environments. Further, the approachability and relative tractability of the fungal hosts we identify makes for hopeful new model systems for deep probes into large virus-host interactions.

Figures


Tree scale: 0.1
NCVOGs Sharing Same Contig
NCVOG0023 (D5-like helicase primase)
NCVOG0038 (DNA pol B)
NCVOG0076 (Superfamily II helicase)
NCVOG0249 (A32-like packaging ATPase)
NCVOG0052 (disulfide (thiol) oxidoreductase)
NCVOG1088 (RNA ligase)
NCVOG1117 (mRNA capping enzyme)
NCVOG1127 (transcription initiation factor IIB)
NCVOG1192 (YqaJ viral recombinase family)
Paramecium bursaria Chlorella virus 1
Cafeteria roenbergensis virus BV-PW1
Ostreococcus tauri virus 1
Kaumoebavirus
Kaumoebavirus
African swine fever virus
Pithovirus sibericum
Cedratvirus A11
Trichoplusia ni ascovirus 2c
Trichoplusia ni ascovirus 2c
Spodoptera frugiperda ascovirus 1a
Heliothis virescens ascovirus 3e
Diadromus pulchellus ascovirus 4a
Ambystoma tigrinum virus
Frog virus 3
Invertebrate iridescent virus 3
Invertebrate iridovirus 22
Invertebrate iridovirus 25
Marseillevirus marseillevirus
Melbournevirus
Lausannevirus
Tunisvirus fontaine2
Mollivirus sibericum
Chrysochromulina ericina virus
Chrysochromulina ericina virus
Phaeocystis globosa virus
Chrysochromulina ericina virus
Cafeteria roenbergensis virus BV-PW1
Ostreococcus tauri virus 1
Ostreococcus tauri virus 1
Cafeteria roenbergensis virus BV-PW1
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Rhizophlyctis rosea JEL0764
Allomyces javanicus California 12
Allomyces arbuscula Burma1F
Allomyces javanicus California 12
Gaertneriomyces semiglobifer Barr 43
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus ARG121
Chytriomyces hyalinus JEL0632
Ectocarpus siliculosus virus 1
Feldmannia species virus
Emiliania huxleyi virus 86
Gonapodya prolifera
Heterosigma akashiwo virus 01
Paramecium bursaria Chlorella virus 1
Ostreococcus tauri virus 1
Ostreococcus tauri virus 1
Ostreococcus tauri virus 1
Ostreococcus tauri virus 1
Ostreococcus tauri virus 1
Paramecium bursaria Chlorella virus 1
Paramecium bursaria Chlorella virus 1
Paramecium bursaria Chlorella virus 1
Paramecium bursaria Chlorella virus 1
Phaeocystis globosa virus
Chrysochromulina ericina virus
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Cafeteria roenbergensis virus BV-PW1
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Cafeteria roenbergensis virus BV-PW1
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Fimicolochytrium jonesii JEL0569
Podochytrium sp. JEL0797
Gonapodya prolifera
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Polychytrium aggegatum JEL0109
Polychytrium aggegatum JEL0109
Polychytrium aggegatum JEL0109
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Catenaria anguillulae PL171
Allomyces javanicus California 12
Allomyces javanicus California 12
Allomyces arbuscula Burma1F
Allomyces javanicus California 12
Allomyces javanicus California 12
Siphonaria sp. JEL0065
Siphonaria sp. JEL0065
Chytriomyces hyalinus JEL0345
Acanthamoeba polyphaga moumouvirus
Siphonaria sp. JEL0065
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell

Figure 4.1 Maximum-likelihood tree of NCVOG0022, the NCVDV Major Capsid Protein. Blue circles indicate nodes with at least 50% bootstrap support.

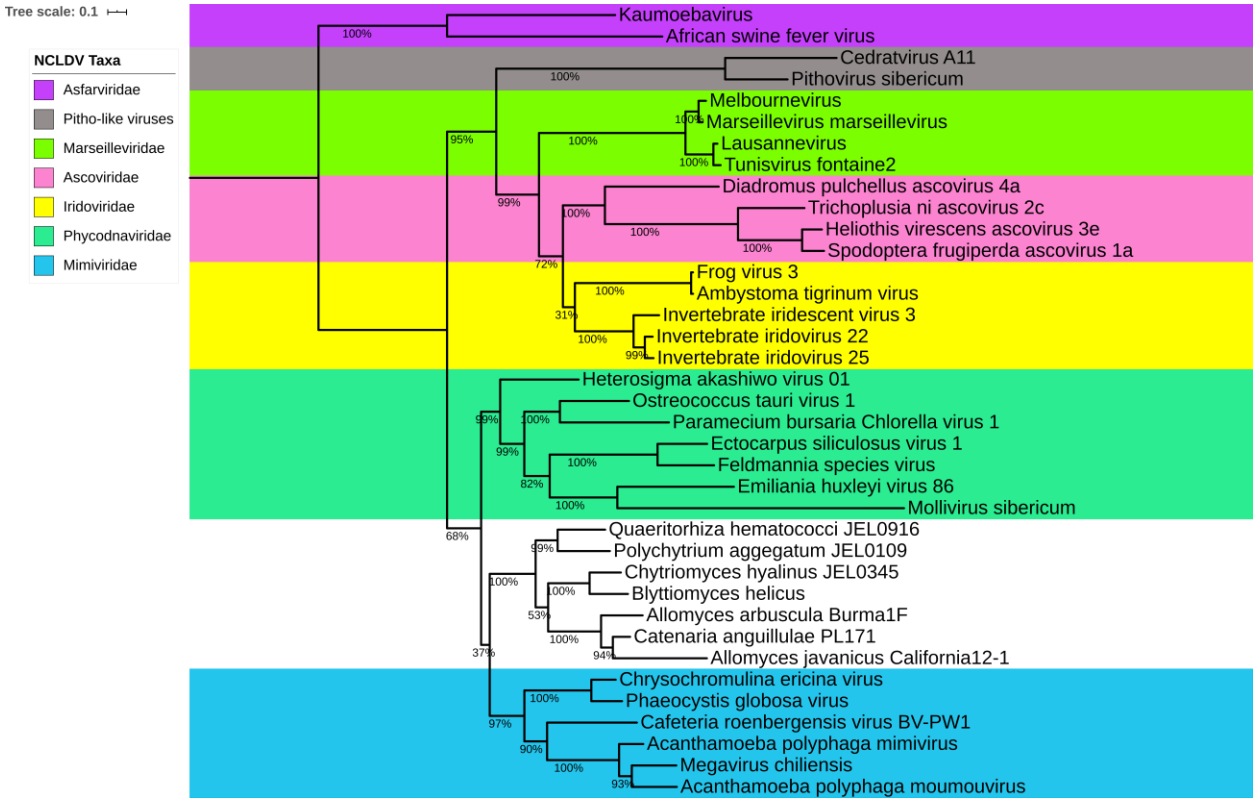


Figure 4.2 Maximum-likelihood tree of the major groups of NCLDVs, including the seven fungal NCLDVs with at least 4/5 core genes (proposed family *Mycodnaviridae*). These five conserved genes (NCVOG0022, NCVOG0023, NCVOG0038, NCVOG0076, NCVOG0249) were concatenated in this phylogenetic reconstruction.

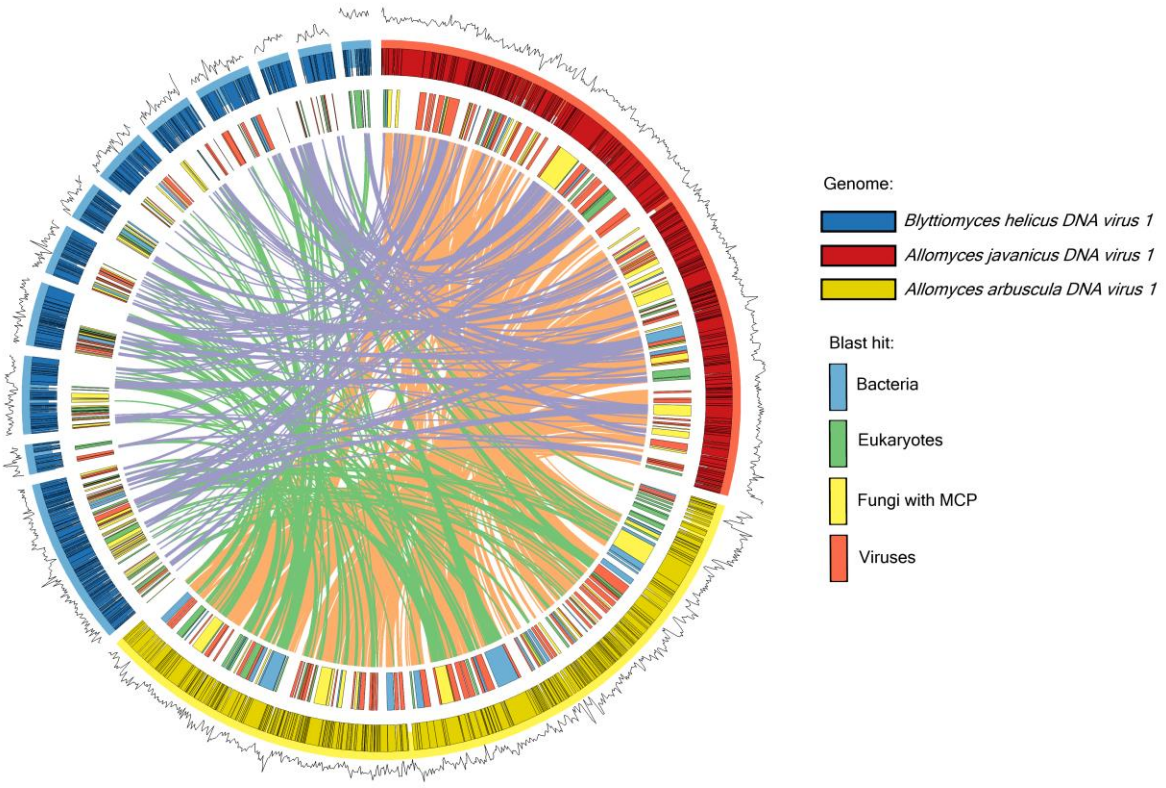


Figure 4.3 Circos plot of three representative *Mycodnaviridae* genomes, *Blyttiomyces helicus DNA virus 1* (blue), *Allomyces javanicus DNA virus 1* (red), and *Allomyces arbuscula DNA virus 1* (yellow). Outermost track shows percent GC across a 500bp sliding window. Next track (blue, red, or yellow) shows the gene density of each contig included in the virome. Innermost track shows genes colored by blast hit to viruses (red), bacteria (blue), eukaryotes (green), or only to fungi shown in this study to have NCLDV major cap proteins (yellow). Center bands connect homologous regions of the genomes.

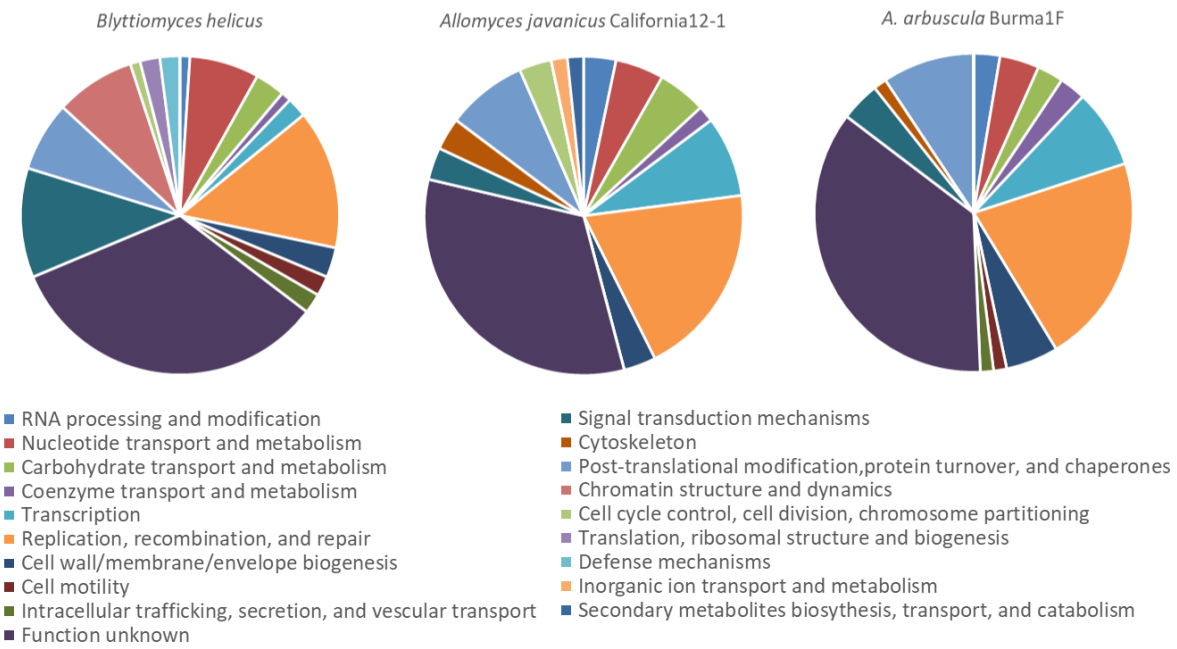


Figure 4.4 Functional annotations of three representative *Mycodnaviridae* genomes. For each genome, the largest portion of orthologous clusters are of unknown function, followed by genes for replication, recombination, and repair.

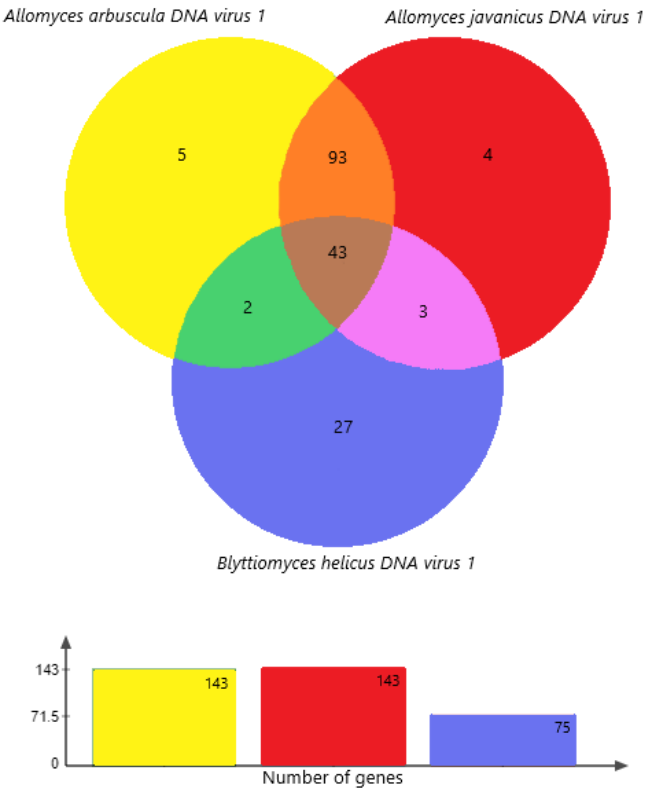


Figure 4.5 Diagram of orthologous gene clusters in *Allomyces arbuscula DNA virus 1* (yellow), *Allomyces javanicus DNA virus 1* (red), and *Blyttiomyces helicus DNA virus 1* (blue), showing 43 orthologous clusters shared by all three.

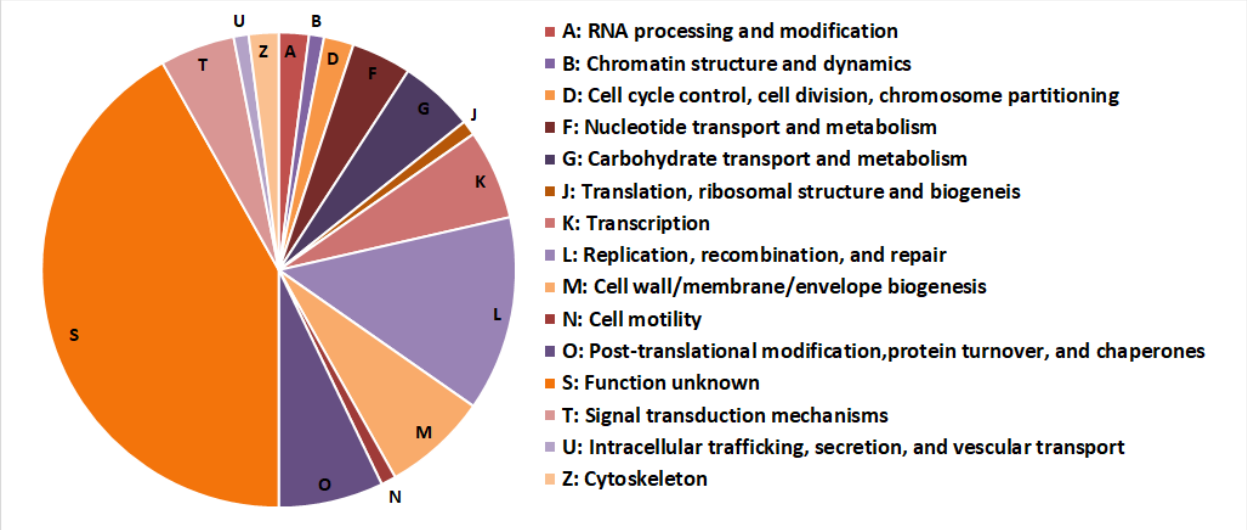


Figure 4.6 Pie chart showing functional assignments of the 43 gene clusters shared by *Blyttiomyces helicus DNA virus 1*, *Allomyces javanicus DNA virus 1*, and *Allomyces arbuscula DNA virus 1*.

Tables

Table 4.1 NCVOG presence in NCLDV families, including proposed family Mycodnaviridae.Table adapted from Yutin & Koonin 2012

		Poxviridae							Asfvariridae
		Goatpox virus Pellor	Swinepox virus	Deerpox virus W-848-83	Variola virus (smallpox virus)	Vaccinia virus	Melanoplus sanguinipes entomopoxvirus	Amsacta moorei entomopoxvirus	African swine fever virus
Functional category	**cluster**								
DNA replication, recombination and repair	NCVOG0004						1	1	2
Virion structure and morphogenesis	NCVOG0022	1	1	1	1	1	1	1	1
DNA replication, recombination and repair	NCVOG0023	1	1	1	1	1	1	1	1
DNA replication, recombination and repair	NCVOG0035						1	1	
DNA replication, recombination and repair	NCVOG0036	1	1	1	1	1	1	1	
DNA replication, recombination and repair	NCVOG0038	1	1	1	1	1	1	1	1
Virion structure and morphogenesis	NCVOG0052	1	1	1	1	1	1	1	1
Uncharacterized	NCVOG0059								1
Transcription and RNA processing	NCVOG0076	1	1	1	1	1	1	1	2
Virion structure and morphogenesis	NCVOG0211	2	2	2	2	2	2	2	1
Virion structure and morphogenesis	NCVOG0249	1	1	1	1	1	1	1	1
Virion structure and morphogenesis	NCVOG0256	1	1	1	1	1	1	1	
Transcription and RNA processing	NCVOG0267	1	1	1	1	1	1	1	1
DNA replication, recombination and repair	NCVOG1060	1	1	1	1	1	1	1	
Transcription and RNA processing	NCVOG1088								1
DNA replication, recombination and repair	NCVOG1115	1	1	1	1	1	1	1	
Transcription and RNA processing	NCVOG1117	1	1	1	1	1	1	1	1
Virion structure and morphogenesis	NCVOG1122	3	3	3	3	3	3	3	
Transcription and RNA processing	NCVOG1127								1
DNA replication, recombination and repair	NCVOG1192							1	1

Phycodnaviridae								Mimiviridae		Ascoviridae			Iridoviridae							Marseillevirus
Paramecium bursaria chlorella virus MT325	Paramecium bursaria Chlorella virus FR483	Acanthocystis turfacea Chlorella virus 1	Paramecium bursaria Chlorella virus 1	Ostreococcus virus OsV5	Feldmannia species virus	Ectocarpus siliculosus virus 1	Emiliania huxleyi virus 86	Mimivirus	Mamavirus	Heliothis virescens ascovirus 3e	Spodoptera frugiperda ascovirus 1a	Trichoplusia ni ascovirus 2c	Invertebrate iridescent virus 6	Aedes taeniorhynchus iridescent virus	Ambystoma tigrinum virus	Frog virus 3	Lymphocystis disease virus 1	Lymphocystis disease virus - isolate China	Infectious spleen and kidney necrosis virus	Marseillevirus
								2	2											1
5	6	5	6	8	1	1	1	4	4	1	1	2	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
								1	1				1	1						
				1				1	1											
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1		1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1
								4	4											
1	1	1	1	1		1	1	1	1				1	1	1	1				1
								1	1	1	2	2	2	1	1	1	1	1	1	
1	1	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1
								3	3											
								2	2											
							1	1	1	1	1	1	1	1	1	1	1	1	1	1
										1	1	2	1	1	1	1	1	1	1	1
								1	1											1
1	1	2	1	1			1	1	1										1	1
								2	2		1	1	1	1	1	1	1	1		
1	1	1	1	1				1	1											1
2	2	2	2	1	1	1		1	1											

Mycodnaviridae (proposed)																
Allomyces javonicus DNA virus 1	Allomyces arbuscula DNA virus 1	Rhizophlyctis rosea JEL0764	Quaeritorhiza hematococci DNA virus 1	Chytriomyces hyalinus DNA virus 1	Gonapodya sp. JEL0774	Chytriomyces hyalinus ARG121	Podochytrium JEL797	Siphonaria sp. 0065	Gaertneriomyces semiglobifer Barr 43	Blyttiomyces helicus DNA virus 1	Catenaria anguillulae DNA virus 1	Chytriomyces hyalinus JEL0632	Gonapodya prolifera	Fimicolochytrium jonesii JEL0569	Polychytrium aggegatum DNA virus 1	**annotation**
																AP (apurinic) endonuclease family 2 - bacterial
1	2	1	8	1	5	1	1	3	1	8	1	1	2	1	3	NCLDV major capsid protein
2	2										1				1	D5-like helicase-primase
																iMV envelope protein p35
																DNA topoisomerase I
1	1	1	1	1	3		1	1		2	2	1			2	DNA polymerase elongation subunit family B
1	1	1	2							1					2	disulfide (thiol) oxidoreductase; Erv1 / Alr family
		3		1						2						FtsJ-like methyltransferase family proteins
4	2			1				1		1	1		1			DNA or RNA helicases of superfamily II (COG1061)
			1													myristylated IMV envelope protein
1	1		1	1	3					2					1	A32-like packaging ATPase
																IMV envelope protein p35
																RNA-helicase DExH-NPH-II
																FLAP-like endonuclease XPG (cd00128)
1			1	1	2		1	1		4		2	1		6	RNA ligase (conserved in irido-, asfa- asco- and Marseille viruses)
																uracil-DNA glycosylase
1	1		2	1	2			3		2			1			mRNA capping enzyme large subunit
																Myristylated protein; pfam03003, DUF230
1	1			1						3					5	transcription initiation factor IIB
1	2		1	1			1	2		4		1			1	YqaJ viral recombinase family

References

Koonin EV & N Yutin. 2010. Origin and evolution of eukaryotic large nucleo-cytoplasmic DNA viruses. *Intervirology.* DOI: 10.1159/000312913

Iyer LM, L Aravind, & EV Koonin. 2001. Common origin of four diverse families of large eukaryotic DNA viruses. *Journal of Virology.* 75(23): 11720-11734

Galindo I & C Alonso. 2017. African swine fever virus: a review. *Viruses.* 9(103) doi:10.3390/v9050103

Wilson WH, JL Etten, & MJ Allen. 2009. The *Phycodnaviridae*: the story of how tiny giants rule the world. *Curr Top Microbiol Immunol.* 328:1-42

Laber CP, Hunter JE, Carvalho F *et al.* 2018. *Coccolithovirus* facilitation of carbon export in the North Atlantic. *Nat Microbiol* 3, 537–547 https://doi.org/10.1038/s41564-018-0128-4

Condit RC. 2007. Vaccinia, Inc.—probing the functional substructure of poxviral replication factories. *Cell Host & Microbe.* 2: 205-207

Forterre P & M Gaïa. 2016. Giant viruses and the origin of modern eukaryotes. *Current Opinion in Microbiology.* 31: 44-49

Guglielmini J, AC Woo, M Krupovic, P Forterre, & M Gaïa. 2019. Diversification of giant and large eukaryotic dsDNA viruses predated the origin of modern eukaryotes. *Proc Natl Acad Sci.* 116(39): 19585- 19592

La Scola B, S Audic, C Robert, L Jungang, X de Lamballerie, M Drancourt, R Birtles, JM Claverie, D Raoult. 2003. A giant virus in amoebae. *Science.* 299: 2033

Rauoult D, S Audic, C Robert, C Abergel, P Renesto, H Ogata, B La Scola, M Suzan, & JM Claverie. 2004.

Boyer M, Madoui M-A, Gimenez G, La Scola B, Raoult D. 2010. Phylogenetic and phyletic studies of informational genes in genomes highlight existence of a 4 domain of life including giant viruses. *PLoS One.* 5:e15530.

Colson P, Gimenez G, Boyer M, Fournous G, Raoult D. 2011. The giant cafeteria roenbergensis virus that infects a widespread marine phagocytic protist is a new member of the fourth domain of life. *PLoS One.* 6:e18935.

Legendre M, Arslan D, Abergel C, & Claverie J. 2012. Genomics of Megavirus and the elusive fourth domain of life. *Comm Int Biol*. 5:102-106.

Yutin N, Wolf YI, & Koonin EV. 2014. Origin of giant viruses from smaller DNA viruses not from a fourth domain of cellular life. *Virology*. 466–467:38-52.

Nasir A, Kim KM, & Caetano-Anolles G. 2012 Giant viruses coexisted with the cellular ancestors and represent a distinct supergroup along with superkingdoms Archaea, Bacteria and Eukarya. *BMC Evol Biol*. 12:156.

Williams TA, Embley TM, & Heinz E. 2011. Informational gene phylogenies do not support a fourth domain of life for nucleocytoplasmic large DNA viruses. *PLoS One.* 6:e21080.

Moreira D, López-García P. 2015. Evolution of viruses and cells: do we need a fourth domain of life to explain the origin of eukaryotes? *Philos Trans R Soc B.* 370:20140327.

Boughalmi M, Saadi H, Pagnier I, Colson P, Fournous G, Raoult D, & La Scola B. 2013 High-throughput isolation of giant viruses of the Mimiviridae and Marseilleviridae families in the Tunisian environment. *Env Microbiol.* 15:2000-2007.

Koonin E, Dolja V, Krupovic M, Varsani A, Wolf Y, Yutin N, Zerbini F, & Kuhn J. 2019. Create a megataxonomic framework, filling all principal taxonomic ranks, for DNA viruses encoding vertical jelly roll-type major capsid proteins. 10.13140/RG.2.2.14886.47684.

Filée J. 2013. Route of NCLDV evolution: the genomic accordion. *Current Opinion in Virology.* 3:595-599

Koonin EV, Krupovic M, & Yutin N. 2015. Evolution of double-stranded DNA viruses of eukaryotes: from bacteriophages to transposons to giant viruses. *Ann. N.Y. Acad. Sci. 1341 (2015) 10–24. doi: 10.1111/nyas.12728*

Filée J, N Pouget, & M Chandler. 2008. Phylogenetic evidence for extensive lateral acquisition of cellular genes by nucleocytoplasmic large DNA viruses. *BMC Evol Biol*. 8:320.

Suhre K. 2005. Gene and genome duplication in Acanthamoeba polyphaga Mimivirus. *J Virol*. 79:14095-14101.

Moniruzzaman M, CA Martinez-Gutierrez, AR Weinheimer, & FO Aylward. 2020. Dynamic genome evolution and complex virocell metabolism of globally-distributed giant viruses. *Nature Communications.* 11: 1710

Iyer LM, S Balaji, EV Koonin, & L Aravind. 2006. Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res*. 117, 156–184

Hingamp, P. *et al*. 2013. Exploring nucleo-cytoplasmic large DNA viruses in Tara oceans microbial metagenomes. *ISME J*. 7, 1678–1695

Moniruzzaman M, LL Wurch, H Alexander, ST Dyhrman, CJ Gobler, & SW Wilhelm. 2017. Virus-host relationships of marine single-celled eukaryotes resolved from metatranscriptomics. *Nature Communications.* 8: 16054

Schulz F. *et al*. 2020. Giant virus diversity and host interactions through global metagenomics. *Nature*. https://doi.org/10.1038/s41586-020-1957-x

Mihara T. *et al*. 2018. Taxon richness of 'Megaviridae' exceeds those of bacteria and archaea in the ocean. *Microbes Environ*. 33: 162–171

Schulz F. *et al.* 2018. Hidden diversity of soil giant viruses. *Nature Communications.* doi: 10.1038/s41467-018-07335-2

Fischer, MG. 2016. Giant viruses come of age. *Current Opinion in Microbiology*. **31**, 50-57

van de Vossenberg BTLH, Warris S, Nguyen HDT, *et al*. 2019. Comparative genomics of chytrid fungi reveal insights into the obligate biotrophic and pathogenic lifestyle of Synchytrium endobioticum. *Sci Rep*. 9(1):8672. doi:10.1038/s41598-019-45128-9

Chang Y, Wang S, Sekimoto S, *et al*. 2015. Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. *Genome Biol Evol*. 7(6):1590-1601. doi:10.1093/gbe/evv090

Ahrendt SR, Quandt CA, Ciobanu D, et al. 2018. Leveraging single-cell genomics to expand the fungal tree of life. *Nat Microbiol*. 3(12):1417-1428. doi:10.1038/s41564-018-0261-0

Mondo SJ, Dannebaum RO, Kuo RC, et al. 2017. Widespread adenine N6-methylation of active genes in fungi. *Nat Genet*. 49(6):964-968. doi:10.1038/ng.3859

Simmons DR, AE Bonds, BT Castillo, RA Clemons, AD Glasco, JM Myers, N Thapa, PM Letcher, MJ Powell, JE Longcore, & TY James. 2020. The collection of zoosporic eufungi at the University of Michigan (CZEUM): introducing a new repository of fully barcoded Chytridiomyceta and Blastocladiomycota cultures. *IMA Fungus.* In press.

James TY, Stenlid J, Olson A, Johannesson H. 2008. Evolutionary significance of imbalanced nuclear ratios within heterokaryons of the basidiomycete fungus Heterobasidion parviporum. *International Journal of Organic Evolution*. 62(9):2279-2296. DOI: 10.1111/j.1558-5646.2008.00462.x.

Nurk S. et al. 2013. Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads. In: Deng M., Jiang R., Sun F., Zhang X. (eds) Research in Computational Molecular Biology. RECOMB 2013. Lecture Notes in Computer Science, vol 7821. Springer, Berlin, Heidelberg

Safonova Y., Bankevich A., Pevzner P.A. (2014) dipSPAdes: Assembler for Highly Polymorphic Diploid Genomes. In: Sharan R. (eds) Research in Computational Molecular Biology. RECOMB 2014. Lecture Notes in Computer Science, vol 8394. Springer, Cham

Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC bioinformatics*, *11*, 119. https://doi.org/10.1186/1471-2105-11-119

Eddy, "A new generation of homology search tools based on probabilistic inference". Genome Informatics, 2009.

Yutin N, Wolf YI, Raoult D, & Koonin EV. 2009. Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution.*Virol. J.* 6, 223.

Katoh, K., & Standley, D.M. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Molecular Biology and Evolution, (30)4. 772–780. doi: 10.1093/molbev/mst010.

Capella-Gutiérrez, S., Silla-Martínez, J.M., & Gabaldón, T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics, (25)15. 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Price, M.N., Dehal, P.S., and Arkin, A.P. (2009) FastTree: Computing Large Minimum-Evolution Trees with Profiles instead of a Distance Matrix. Molecular Biology and Evolution 26:1641-1650, doi:10.1093/molbev/msp077.

Kang DD, F Li, E Kirton, *et al*. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. doi:10.7717/peerj.7359

Darriba, D., Taboada, G.L., Doallo, R., & Posada, D. 2011. ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics, (27)8. 1164–1165. doi: 10.1093/bioinformatics/btr088

Stamatakis, A. 2014. RaxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics, (30)9. 1312–1313. doi: 10.1093/bioinformatics/btu033

Krzywinski M, *et al.* 2009. Circos: an Information Aesthetic for Comparative Genomics. *Genome Res* 19:1639-1645

Xu L, Z Dong, L Fang, *et al*. 2019. OrthoVenn2: a web server for whole-genome comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res*. 47(W1):W52-W58. doi:10.1093/nar/gkz333

Huerta-Cepas J, Forslund K, Coelho LP, et al. 2017. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol Biol Evol*. 34(8):2115-2122. doi:10.1093/molbev/msx148

Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund S, Cook H, Mende D, Letunic I, Rattei T, Jensen L, von Mering C, & P Bork. 2019. *Nucleic Acids Res*. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Jan 8; 47(Database issue): D309–D314. doi: 10.1093/nar/gky1085

Hoffmann R, R Michel, K-D Müller, & E Schmid. 1998. Archaea-like endocytobiotic organisms isolated from Acanthamoeba sp. (Gr II). Endocytobiosis Cell Res, 12:185-188.

Scheid P, S Pressmar, G Richard, R Zöller, & R Michel. 2008. An extraordinary endocytobiont in Acanthamoeba sp. isolated from a patient with keratitis. Parasitol Res, 102:945-950.

Scheid, P. 2016. A strange endocytobiont revealed as largest virus. *Current Opinion in Microbiology.* 31: 58-62

Cantino EC & EA Horenstein. 1956. Gamma and the cytoplasmic control of differentiation in *Blastocladiella. Mycologia.* 48:3, 443-446.

Cantino EC & JP Mack. 1969. Nova Hedwigia. 18, 115.

Myers RB & EC Cantino. 1971. DNA profile of the spore of *Blastocladiella emersonii*: evidence for γ particle DNA. *Arch. Mikrobiol*. 78: 252-267

Matsumae, A, RB Myers, & EC Cantino. 1970. Comparative numbers of γ particles in the flagellate cells of various species and mutants of *Blastocladiella*. *J. Gen. Appl. Microbiol*. 16: 443-453

Bartow WE. 1979. Ultrastructure of the formation of γ particles during zoosporogenesis in *Allomyces macrogynus* and *Catenaria anguillulae. Experimental Mycology.* 3: 28-41

Barstow WE & JS Lovett. 1975. Formation of γ particles during zoosporogenesis in *Blastocladiella emersonii*. *Mycologia*, 67(3): 518-529.

Romero-Brey I & R Bartenschlager. 2016. Endoplasmic reticulum: the favorite intracellular niche for viral replication and assembly. *Viruses.* 8:160; doi: 10.3390/v8060160

Lessie, P. E., and J. S. Lovett. 1968. Ultrastructural changes during sporangium formation and zoospore differentiation in Blastocladiella emersonii. Am. J. Bot. 55:220-236.

Truesdell, L. C., and E. C. Cantino. 1970. Decay of gamma particles in germinating zoospores of Blastocladiella emersonii. Arch. Mikrobiol. 70:378-392.

Li P, S Wang, L Zhang, D Qiu, X Zhou, & L Guo. 2020. A tripartite ssDNA mycovirus from a plant pathogenic fungus is infectious as cloned DNA and purified virions. *Science Advances.* 6(14) DOI: 10.1126/sciadv.aay9634

Yu X, B Li, Y Fu, D Jiang, SA Ghabrial, G Li, Y Peng, J Xie, J Cheng, J Huang, X Yi. 2010. A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus. *Proc. Natl. Acad. Sci*. 107: 8387-8392

Moniruzzaman M, ER Gann, & SW Wilhelm. 2018. Infection by a giant virus (AaV) induces widespread physiological reprogramming in *Aureococcus anophagefferens* CCMP1984- a harmful bloom algae. *Frontiers in Microbiology.* 9:752 doi: 10.3389/fmicb.2018.00752

Brussaard CP, SW Wilhelm, F Thingstad, MG Weinbauer,G Bratbak, M Heldal, *et al*. 2008. Global-scale processes with a nanoscale drive: the role of marine viruses. *ISME J*. 2, 575–578. doi: 10.1038/ismej.2008.31

Supplement

Supplemental Table 4.1 Isolates screened for evidence of NCLDVs, their project information, and the technology used to sequence their genome.

SPECIES	GENUS	SPECIES	STRAIN	Project	tech
Allomyces_arbuscula_Burma_1F.LCG	Allomyces	arbuscula	Burma_1F	1978	illumina
Allomyces_javanicus_California_12.LCG	Allomyces	javanicus	California_12	1978	illumina
Allomyces_macrogynus_ATCC_38327	Allomyces	macrogynus	ATCC_38327	JGI	sanger
Anaeromyces_robustus_v1.0	Anaeromyces	robustus	S4	JGI	NA
Batrachochytrium_dendrobatidis_JAM81_v1.0	Batrachochytrium	dendrobatidis	JAM81	JGI	sanger
Batrachochytrium_dendrobatidis_JEL423	Batrachochytrium	dendrobatidis	JEL423	NA	NA
Batrachochytrium_salamandrivorans_BS	Batrachochytrium	salamandrivorans	BS	NA	NA
Blastocladiella_britannica_v1.0	Blastocladiella	britannica	JEL711	JGI	NA
Blyttiomyces_helicus_single-cell_v1.0	Blyttiomyces	helicus	NA	JGI	NA
Blyttiomyces_sp._JEL0837.LCG	Blyttiomyces	sp.	JEL0837	1978	illumina
Boothiomyces_macroporosum_PLAUS21.LCG	Boothiomyces	macroporosum	PLAUS21	UMICH	NA
Boothiomyces_sp._JEL0838.LCG	Boothiomyces	sp.	JEL0838	1978	illumina
Boothiomyces_sp._JEL0866.LCG	Boothiomyces	sp.	JEL0866	1978	illumina
Borealophlyctis_nickersoniae_WJD170.LCG	Borealophlyctis	nickersoniae	WJD170	UMICH	NA
Catenaria_anguillulae_PL171_v2.0	Catenaria	anguillulae	PL171	JGI	NA
Caulochytrium_protostelioides_ATCC_52028_v1.0	Caulochytrium	protostelioides	ATCC_52028	JGI	NA
Chytridium_lagenaria_Arg66_v1.0	Chytridium	lagenaria	Arg66	JGI	pacbio
Chytriomyces_confervae_CBS_675.73	Chytriomyces	confervae	CBS_675.73	Syn	NA
Chytriomyces_hyalinus_ARG085.LCG	Chytriomyces	hyalinus	ARG085	1978	illumina
Chytriomyces_hyalinus_ARG121.LCG	Chytriomyces	hyalinus	ARG121	1978	illumina
Chytriomyces_hyalinus_JEL0176.LCG	Chytriomyces	hyalinus	JEL0176	1978	illumina
Chytriomyces_hyalinus_JEL0345.LCG	Chytriomyces	hyalinus	JEL0345	1978	illumina
Chytriomyces_hyalinus_JEL632_v1.0	Chytriomyces	hyalinus	JEL632	JGI	pacbio
Chytriomyces_sp._MP_71_v1.0	Chytriomyces	sp.	MP_71	JGI	NA

Cladochytrium_replicatum_JEL714_v1.0	Cladochytrium	replicatum	JEL714	JGI	NA
Cladochytrium_tenue_CCIBt4013.v0.LCG	Cladochytrium	tenue	CCIBt4013	UMICH	NA
Clydaea_vesicula_JEL0476.LCG	Clydaea	vesicula	JEL0476	UMICH	NA
Coelomomyces_lativittatus_CIRM-AVA-1-Amber.LCG	Coelomomyces	lativittatus	CIRM-AVA-1-Amber	UCR	NA
Coelomomyces_lativittatus_CIRM-AVA-1-Meiospore.LCG	Coelomomyces	lativittatus	CIRM-AVA-1-Meiospore	UCR	NA
Coelomomyces_lativittatus_CIRM-AVA-1-Orange.LCG	Coelomomyces	lativittatus	CIRM-AVA-1-Orange	UCR	NA
Dinochytrium_kinnereticum_KLL_TL_06062013.LCG	Dinochytrium	kinnereticum	KLL_TL_06062013	1978	illumina
Entophlyctis_helioformis_JEL805_v1.0	Entophlyctis	helioformis	JEL805	JGI	pacbio
Entophlyctis_luteolus_JEL0120.LCG	Entophlyctis	luteolus	JEL0120	1978	illumina
Entophlyctis_luteolus_JEL0129.LCG	Entophlyctis	luteolus	JEL0129	1978	illumina
Entophlyctis_sp._JEL0112.LCG	Entophlyctis	sp.	JEL0112	1978	illumina
Fimicolochytrium_jonesii_JEL569_v1.0	Fimicolochytrium	jonesii	JEL569	JGI	pacbio
Gaertneriomyces_semiglobifer_Barr_43_v1.0	Gaertneriomyces	semiglobifer	Barr_43	JGI	NA
Gaertneriomyces_sp._JEL0708.LCG	Gaertneriomyces	sp.	JEL0708	1978	illumina
Geranomyces_michiganensis_JEL0563.LCG	Geranomyces	michiganensis	JEL0563	1978	illumina
Geranomyces_variabilis_JEL0379.LCG	Geranomyces	variabilis	JEL0379	1978	illumina
Geranomyces_variabilis_JEL0389.LCG	Geranomyces	variabilis	JEL0389	1978	illumina
Geranomyces_variabilis_JEL0566.LCG	Geranomyces	variabilis	JEL0566	1978	illumina
Geranomyces_variabilis_JEL0567.LCG	Geranomyces	variabilis	JEL0567	1978	illumina
Geranomyces_variabilis_JEL559_v1.0	Geranomyces	variabilis	JEL559	JGI	NA
Globomyces_pollinis-pini_Arg68_v1.0	Globomyces	pollinis-pini	Arg68	JGI	NA
Gonapodya_prolifera_v1.0	Gonapodya	prolifera	JEL0478	JGI	NA
Gonapodya_sp._JEL0774.LCG	Gonapodya	sp.	JEL0774	1978	illumina
Gorgonomyces_haynaldii_MP57_v1.0	Gorgonomyces	haynaldii	MP57	JGI	pacbio
Homolaphlyctis_polyrhiza_JEL142_v1.0	Homolaphlyctis	polyrhiza	JEL142	JGI	NA
Hyaloraphidium_curvatum_SAG235-1_v1.0	Hyaloraphidium	curvatum	SAG235-1	JGI	pacbio
Irineochytrium_annulatum_JEL0729.LCG	Irineochytrium	annulatum	JEL0729	UMICH	NA
Kappamyces_sp._JEL0680.LCG	Kappamyces	sp.	JEL0680	1978	illumina
Kappamyces_sp._JEL0829.LCG	Kappamyces	sp.	JEL0829	1978	illumina
Lobulomyces_angularis_JEL0522.LCG	Lobulomyces	angularis	JEL0522	1978	illumina
Neocallimastix_californiae_G1_v1.0	Neocallimastix	californiae	G1	JGI	NA
Nowakowskiella_sp._JEL0078.LCG	Nowakowskiella	sp.	JEL0078	UMICH	NA
Nowakowskiella_sp._JEL0407.LCG	Nowakowskiella	sp.	JEL0407	UMICH	NA
Obelidium_mucronatum_JEL802_v1.0	Obelidium	mucronatum	JEL802	JGI	pacbio

Olpidium_bornovanus_UCB_F19785.Olpbor1	Olpidium	bornovanus	UCB_F19785.Olpbor1	JGI	NA
Olpidium_sp._PSC023	Olpidium	sp.	PSC023	UMICH	NA
Paraphysoderma_sedebokerense_JEL821_v1.0	Paraphysoderma	sedebokerense	JEL821	JGI	NA
Pecoramyces_ruminatium_C1A	Pecoramyces	ruminatium	C1A	JGI	NA
Phlyctochytrium_bullatum_JEL0754.LCG	Phlyctochytrium	bullatum	JEL0754	1978	illumina
Phlyctochytrium_planicorne_JEL0388.LCG	Phlyctochytrium	planicorne	JEL0388	1978	illumina
Physocladia_obscura_JEL0513.LCG	Physocladia	obscura	JEL0513	UMICH	NA
Piromyces_finnis_v3.0	Piromyces	finnis	NA	JGI	NA
Piromyces_sp._E2_v1.0	Piromyces	sp.	E2	JGI	NA
Podochytrium_sp._JEL0797.LCG	Podochytrium	sp.	JEL0797	1978	illumina
Polychytrium_aggregatum_JEL109_v1.0	Polychytrium	aggregatum	JEL109	JGI	NA
Powellomyces_hirtus_BR81_v1.0	Powellomyces	hirtus	BR81	JGI	pacbio
Powellomyces_hirtus_CBS_809.83	Powellomyces	hirtus	CBS_809.83	Syn	NA
Quaeritorhiza_haematococci_JEL0916.LCG	Quaeritorhiza	haematococci	JEL0916	UMICH	NA
Rhizoclosmatium_globosum_JEL800_v1.0	Rhizoclosmatium	globosum	JEL800	JGI	NA
Rhizoclosmatium_hyalinum_JEL0917.LCG	Rhizoclosmatium	hyalinum	JEL0917	1978	illumina
Rhizoclosmatium_sp._JEL0117.LCG	Rhizoclosmatium	sp.	JEL0117	1978	illumina
Rhizophlyctis_rosea_JEL0318.LCG	Rhizophlyctis	rosea	JEL0318	UMICH	NA
Rhizophlyctis_rosea_JEL0764.LCG	Rhizophlyctis	rosea	JEL0764	1978	illumina
Rozella_allomycis_CSF55_v1.0	Rozella	allomycis	CSF55	JGI	NA
Rozella_multimorpha	Rozella	multimorpha	NA	UMICH	NA
Rozella_rhizoclosmatii	Rozella	rhizoclosmatii	NA	UMICH	NA
Rozella_sp._PSC023	Rozella	sp.	PSC023	UMICH	NA
Siphonaria_sp._JEL0065.LCG	Siphonaria	sp.	JEL0065	1978	illumina
Spizellomyces_punctatus_DAOM_BR117	Spizellomyces	punctatus	DAOM_BR117	JGI	sanger
Spizellomyces_sp._palustris_CBS_455.65	Spizellomyces	sp.	CBS_455.65	Syn	NA
Synchytrium_endobioticum_MB42	Synchytrium	endobioticum	MB42	Syn	NA
Synchytrium_microbalum_JEL517	Synchytrium	microbalum	JEL517	Syn	NA
Terramyces_sp._JEL0728.LCG	Terramyces	sp.	JEL0728	1978	illumina
Thoreauomyces_humboldtii_JEL0095.LCG	Thoreauomyces	humboldtii	JEL0095	UMICH	NA
Triparticalcar_arcticum_BR59_v1.0	Triparticalcar	arcticum	BR59	JGI	pacbio
Unknown_Chytridiales_sp._JEL0842.LCG	Unknown	Chytridiales	JEL0842	1978	illumina
Unknown_Rhizophydiales_sp._JEL0801.LCG	Unknown	Rhizophydiales	JEL0801	1978	illumina
Unknown_unknown_JEL0888.LCG	Unknown	unknown	JEL0888	UMICH	NA
Zopfochytrium_polystomum_WB228_v1.0	Zopfochytrium	polystomum	WB228	JGI	NA

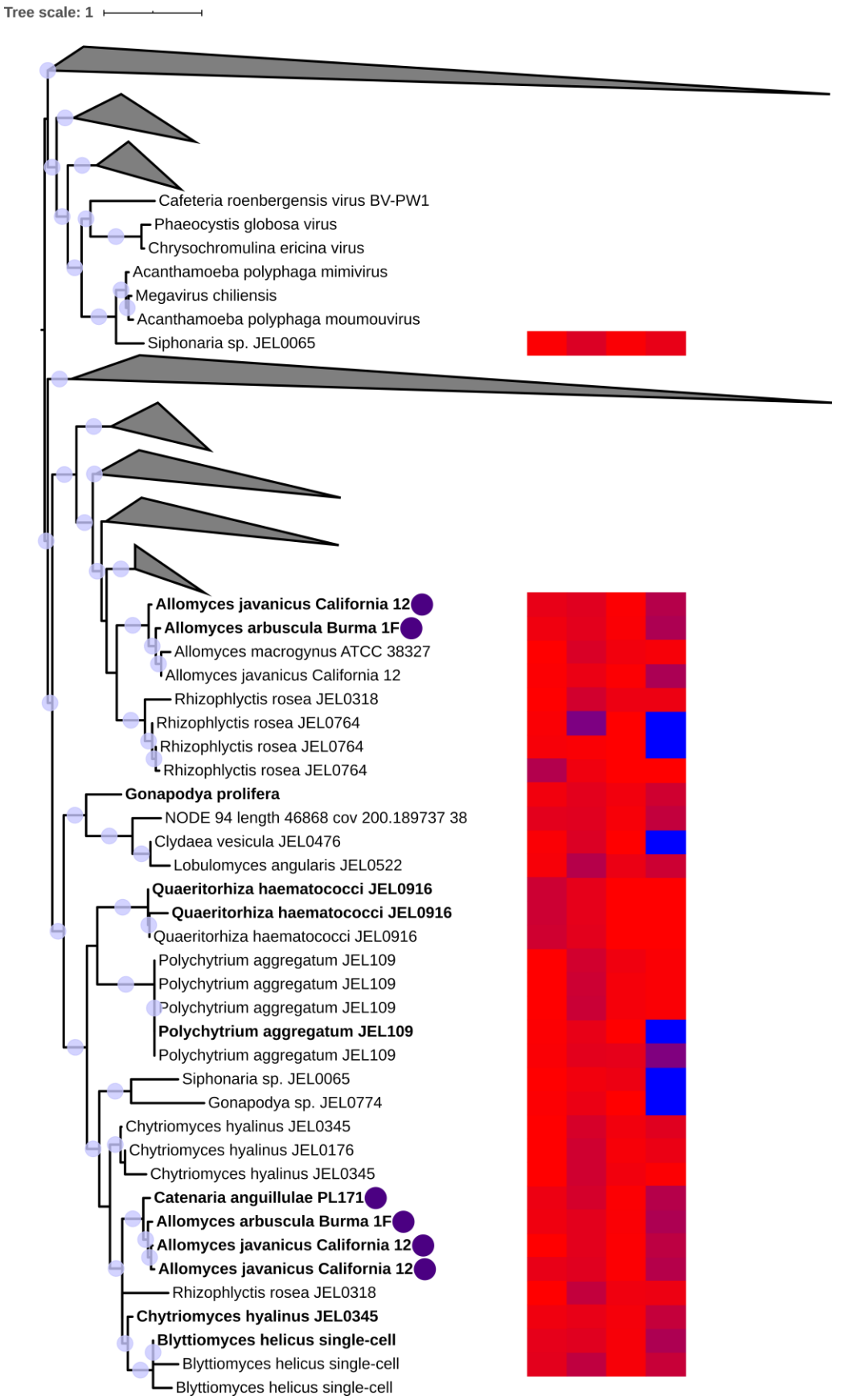
Supplemental Table 4.2 RefSeq accessions of reference NCLDVs.

RefSeq ID	Tree Label	RefSeq Accession
N_uncl_cedvir	Cedratvirus A11	GCF_001995575.1
N_uncl_pitsib	Pithovirus sibericum	GCF_000916835.1
N_uncl_molsib	Mollivirus sibericum	GCF_001292995.1
N_asfa_afswfe	African swine fever virus	GCF_003032865.1
N_uncl_kaumoe	Kaumoebavirus	GCF_002116175.1
N_mars_lausan	Lausannevirus	GCF_000893455.1
N_mars_marsei	Marseillevirus marseillevirus	GCF_000887095.1
N_mars_melbou	Melbournevirus	GCF_000924835.1
N_mars_tunifo	Tunisvirus fontaine2	GCF_002826725.1
N_asco_dipuas	Diadromus puchellus ascovirus 4a	GCF_000881595.1
N_asco_hevias	Heliothis virescens ascovirus 3e	GCF_000871485.1
N_asco_spofru	Spodoptera frugiperda ascovirus 1a	GCF_000867605.1
N_asco_tricni	Trichoplusia ni ascovirus 2c	GCF_000868565.1
N_irid_amtivi	Ambystoma tigrinum virus	GCF_000841005.1
N_irid_frogvi	Frog virus 3	GCF_000844425.1
N_irid_inir22	Invertebrate iridovirus 22	GCF_000909775.1
N_irid_inir25	Invertebrate iridovirus 25	GCF_000914535.1
N_irid_irvi03	Invertebrate iridescent virus 3	GCF_000869125.1
N_phyc_ecsivi	Ectocarpus siliculosus virus	GCF_000839765.1
N_phyc_emhu86	Emiliania huxleyi virus 86	GCF_000865825.1
N_phyc_feldvi	Feldmannia species virus	GCF_000874805.1
N_phyc_osttau	Ostreococcus tauri virus 1	GCF_000885975.1
N_phyc_pabuch	Paramecium bursaria Chlorella virus 1	GCF_000847045.1
N_mimi_acpomi	A. polyphaga mimivirus	GCF_000888735.1
N_mimi_acpomo	A. polyphaga moumouvirus	GCF_000904035.1
N_mimi_carovi	Cafeteria roenbergensis virus	GCF_000889395.1
N_mimi_megchi	Megavirus chiliensis	GCF_000893915.1
N_phyc_chervi	Chrysochromulina ericina virus	GCF_001399245.1
N_phyc_hetaka	Heterosigma akashiwo virus HaV53	GCF_002827745.1

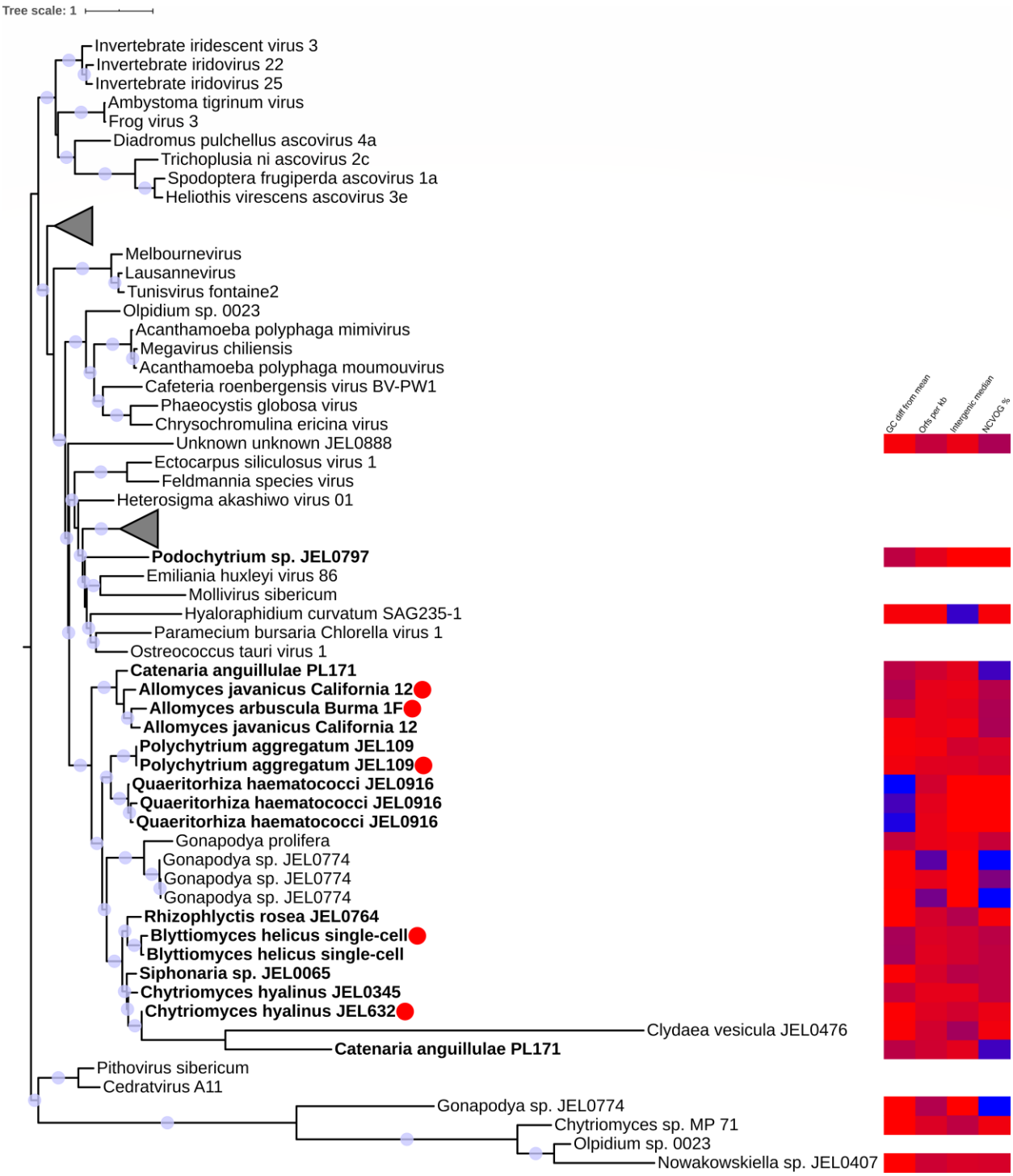
Supplemental Table 4.3 Fungal NCLDV genome sizes per ViralRecall. Asterisk denotes genome which appears to be contaminated by fungal contigs, which was removed from the average per Grubbs' test of outlier detection ($p<0.05$).

Virus	**Sum ViralRecall Contigs (bp)**
Allomyces javanicus DNA virus 1	287,051
Allomyces arbuscula DNA virus 1	346,647
Catenaria anguillulae DNA virus 1	97,938
Blyttiomyces helicus DNA virus 1	301,285
Chytriomyces hyalinus DNA virus 1	251,520
*Polychytrium aggregatum DNA virus 1**	3,539,461
Quaetitorhiza hematococci DNA virus 1	566,424
Mean	308,477.5

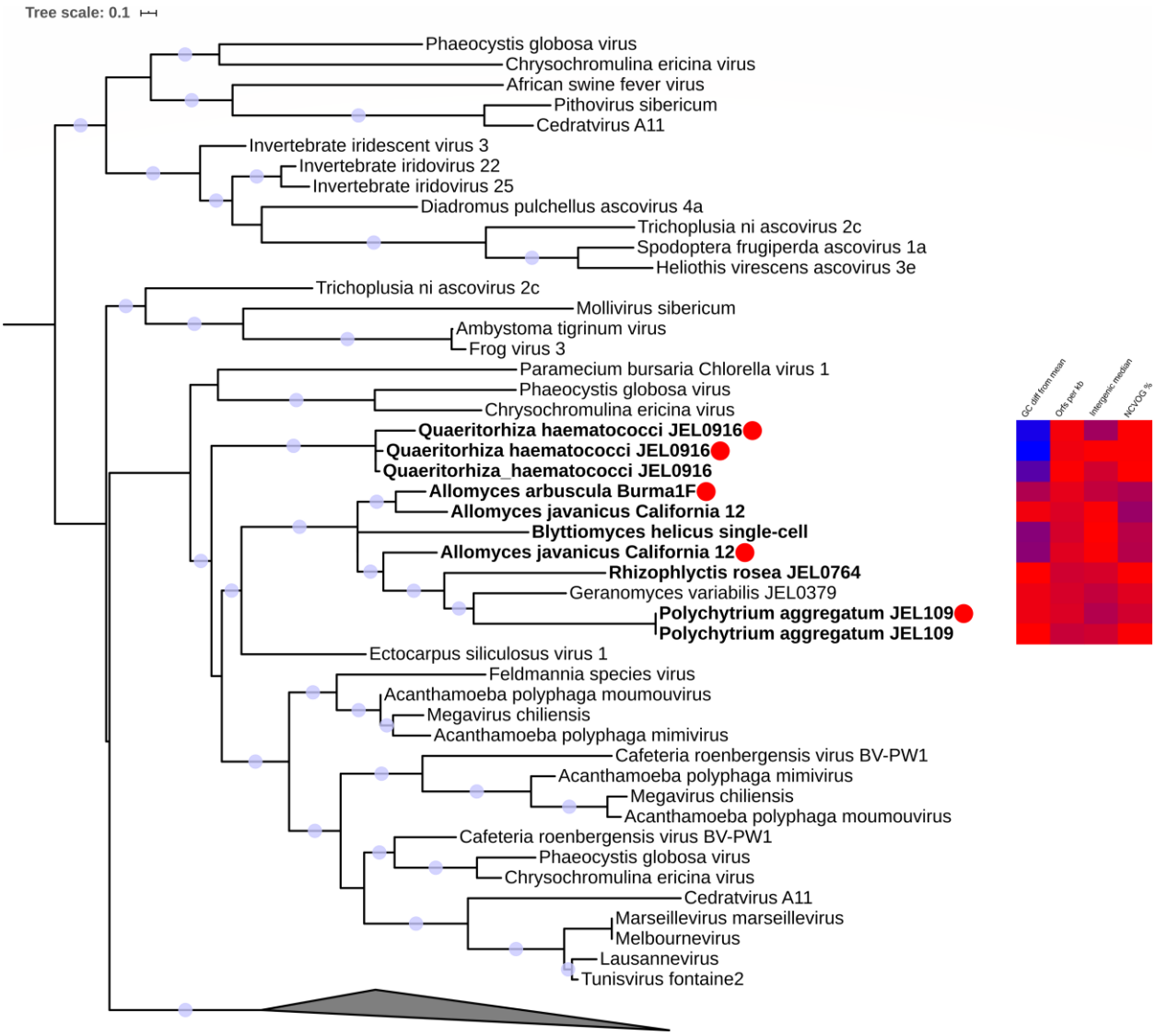
Supplementary Figure 4.1 Gene tree of NCVOG0023 (D5-like helicase-primase). Bolded leaves are fungal NCLDV genes that met our criteria. Purple circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, ORFs per kb sequence, intergenic median, and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).


Tree scale: 1
Cafeteria roenbergensis virus BV-PW1
Phaeocystis globosa virus
Chrysochromulina ericina virus
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Siphonaria sp. JEL0065
Allomyces javanicus California 12
Allomyces arbuscula Burma 1F
Allomyces macrogynus ATCC 38327
Allomyces javanicus California 12
Rhizophlyctis rosea JEL0318
Rhizophlyctis rosea JEL0764
Rhizophlyctis rosea JEL0764
Rhizophlyctis rosea JEL0764
Gonapodya prolifera
NODE 94 length 46868 cov 200.189737 38
Clydaea vesicula JEL0476
Lobulomyces angularis JEL0522
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Siphonaria sp. JEL0065
Gonapodya sp. JEL0774
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus JEL0176
Chytriomyces hyalinus JEL0345
Catenaria anguillulae PL171
Allomyces arbuscula Burma 1F
Allomyces javanicus California 12
Allomyces javanicus California 12
Rhizophlyctis rosea JEL0318
Chytriomyces hyalinus JEL0345
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell

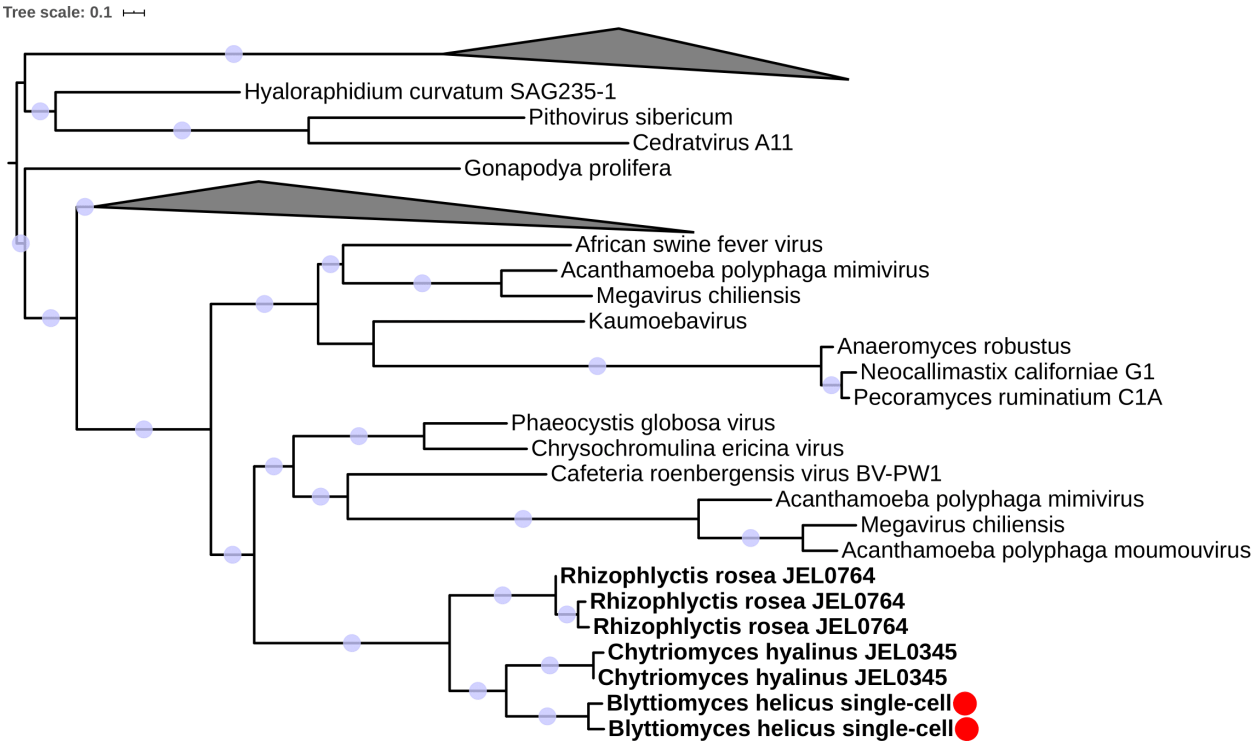
Supplementary Figure 4.2 Gene tree of NCVOG0038 (DNA polymerase elongation subunit family B). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, ORFs per kb sequence, intergenic median, and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).



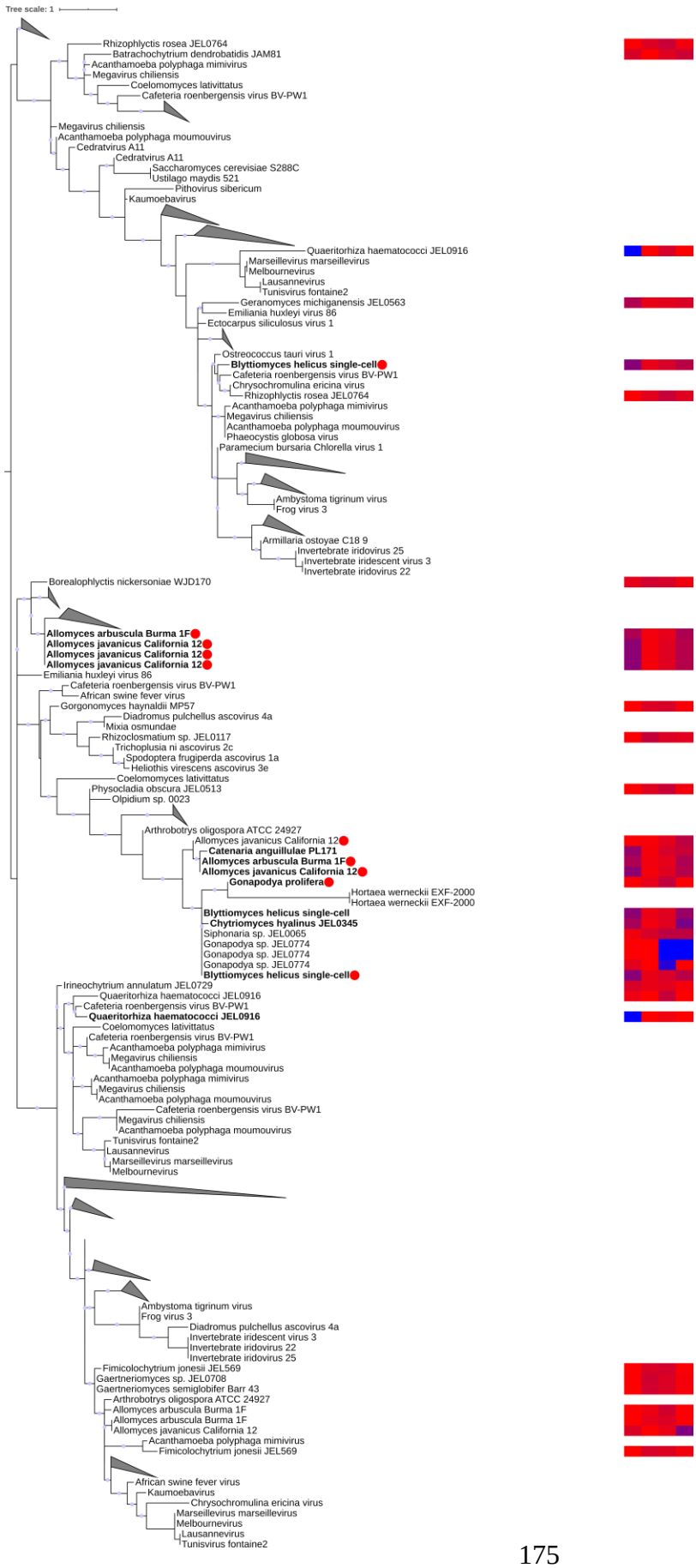
Supplementary Figure 4.3 Gene tree of NCVOG0052 (disulfide (thiol) oxidoreductase; Erv1/Alr family). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, ORFs per kb sequence, intergenic median, and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).



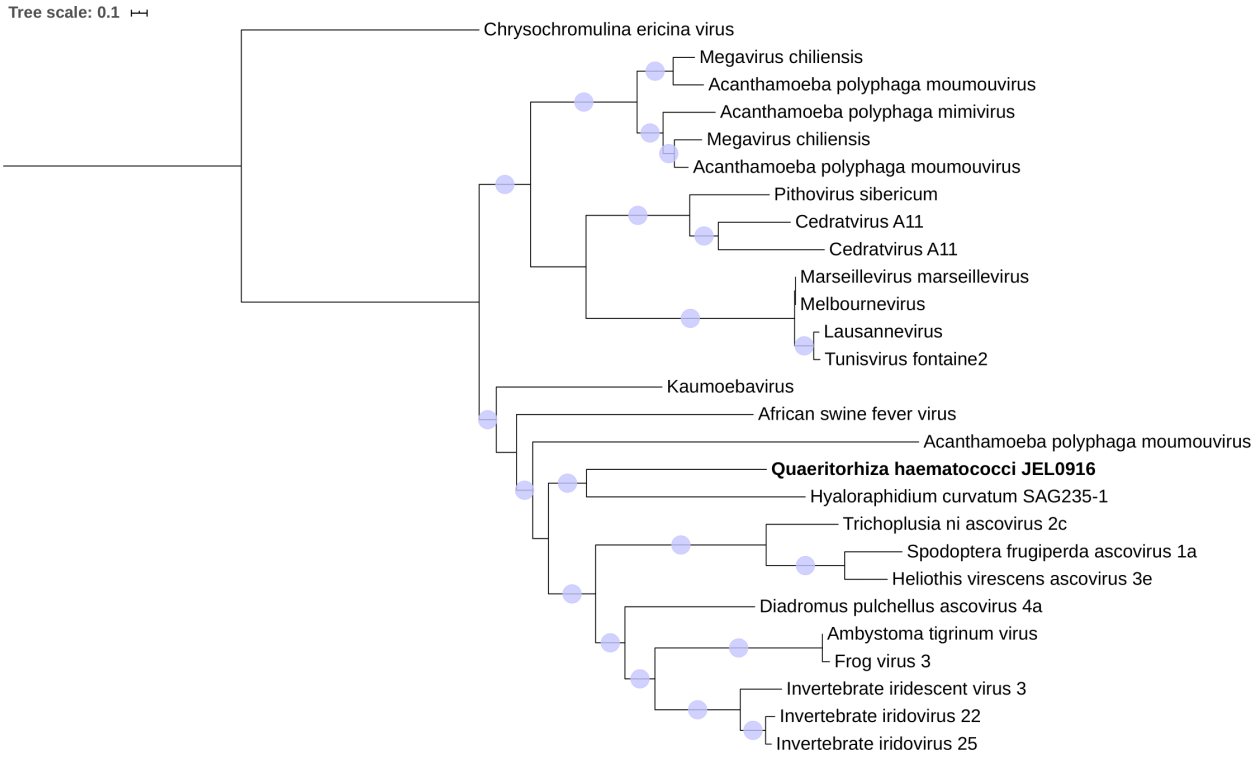
Supplementary Figure 4.4 Gene tree of NCVOG0059 (FtsJ-like methyltransferase family proteins). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot.



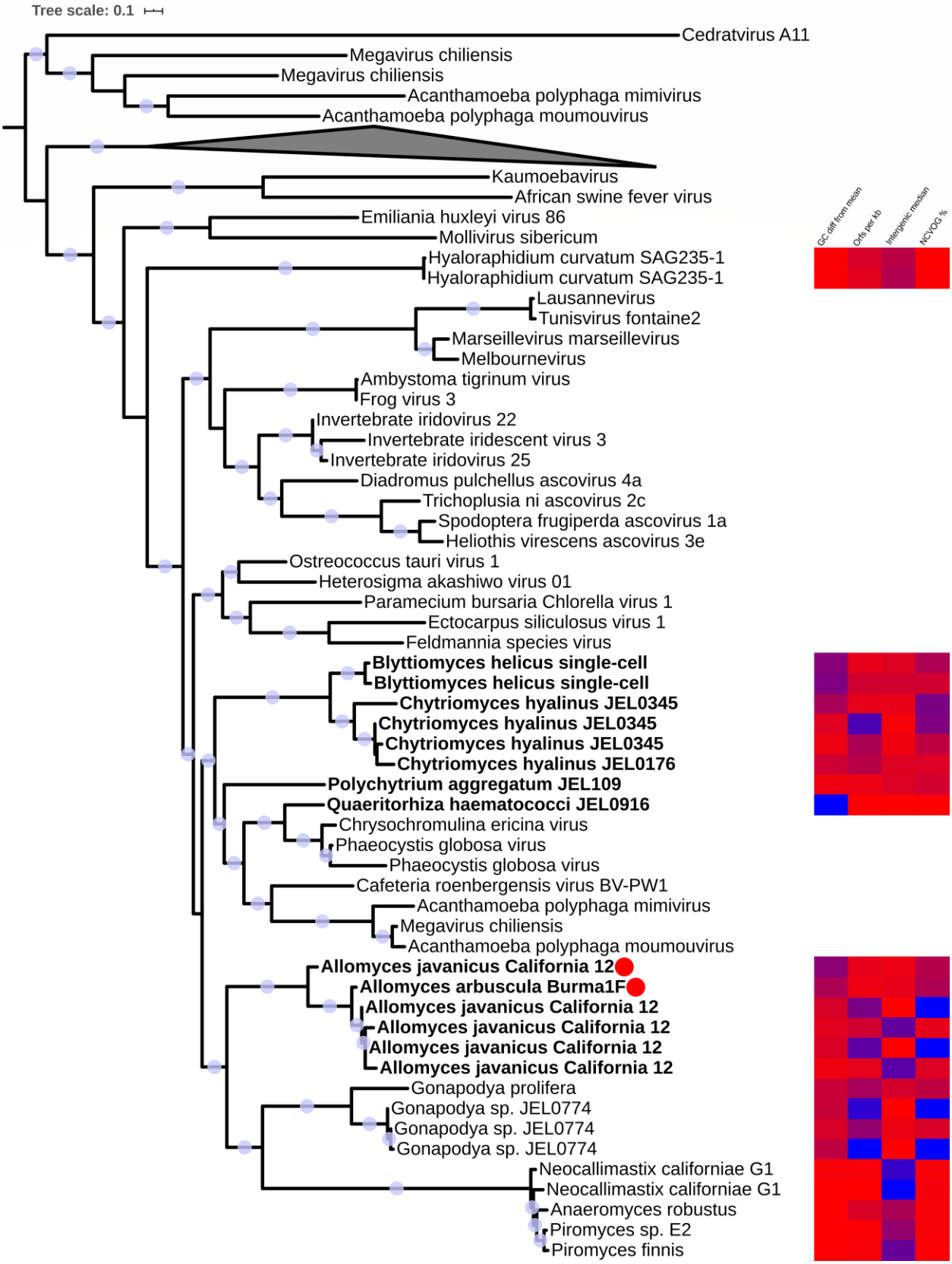
Supplementary Figure 4.5 Gene tree of NCVOG0076 (DNA or RNA helicase of Superfamily II). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, intergenic median, ORFs per kb sequence and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).


Tree scale: 1
Rhizophlyctis rosea JEL0764
Batrachochytrium dendrobatidis JAM81
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Coelomomyces lativittatus
Cafeteria roenbergensis virus BV-PW1
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Cedratvirus A11
Cedratvirus A11
Saccharomyces cerevisiae S288C
Ustilago maydis 521
Pithovirus sibericum
Kaumoebavirus
Quaeritorhiza haematococci JEL0916
Marseillevirus marseillevirus
Melbournevirus
Lausannevirus
Tunisvirus fontaine2
Geranomyces michiganensis JEL0563
Emiliania huxleyi virus 86
Ectocarpus siliculosus virus 1
Ostreococcus tauri virus 1
Blyttiomyces helicus single-cell
Cafeteria roenbergensis virus BV-PW1
Chrysochromulina ericina virus
Rhizophlyctis rosea JEL0764
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Phaeocystis globosa virus
Paramecium bursaria Chlorella virus 1
Ambystoma tigrinum virus
Frog virus 3
Armillaria ostoyae C18 9
Invertebrate iridovirus 25
Invertebrate iridescent virus 3
Invertebrate iridovirus 22
Borealophlyctis nickersoniae WJD170
Allomyces arbuscula Burma 1F
Allomyces javanicus California 12
Allomyces javanicus California 12
Allomyces javanicus California 12
Emiliania huxleyi virus 86
Cafeteria roenbergensis virus BV-PW1
African swine fever virus
Gorgonomyces haynaldii MP57
Diadromus pulchellus ascovirus 4a
Mixia osmundae
Rhizoclosmatium sp. JEL0117
Trichoplusia ni ascovirus 2c
Spodoptera frugiperda ascovirus 1a
Heliothis virescens ascovirus 3e
Coelomomyces lativittatus
Physocladia obscura JEL0513
Olpidium sp. 0023
Arthrobotrys oligospora ATCC 24927
Allomyces javanicus California 12
Catenaria anguillulae PL171
Allomyces arbuscula Burma 1F
Allomyces javanicus California 12
Gonapodya prolifera
Hortaea werneckii EXF-2000
Hortaea werneckii EXF-2000
Blyttiomyces helicus single-cell
Chytriomyces hyalinus JEL0345
Siphonaria sp. JEL0065
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Blyttiomyces helicus single-cell
Irineochytrium annulatum JEL0729
Quaeritorhiza haematococci JEL0916
Cafeteria roenbergensis virus BV-PW1
Quaeritorhiza haematococci JEL0916
Coelomomyces lativittatus
Cafeteria roenbergensis virus BV-PW1
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Cafeteria roenbergensis virus BV-PW1
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Tunisvirus fontaine2
Lausannevirus
Marseillevirus marseillevirus
Melbournevirus
Ambystoma tigrinum virus
Frog virus 3
Diadromus pulchellus ascovirus 4a
Invertebrate iridescent virus 3
Invertebrate iridovirus 22
Invertebrate iridovirus 25
Fimicolochytrium jonesii JEL569
Gaertneriomyces sp. JEL0708
Gaertneriomyces semiglobifer Barr 43
Arthrobotrys oligospora ATCC 24927
Allomyces arbuscula Burma 1F
Allomyces arbuscula Burma 1F
Allomyces javanicus California 12
Acanthamoeba polyphaga mimivirus
Fimicolochytrium jonesii JEL569
African swine fever virus
Kaumoebavirus
Chrysochromulina ericina virus
Marseillevirus marseillevirus
Melbournevirus
Lausannevirus
Tunisvirus fontaine2

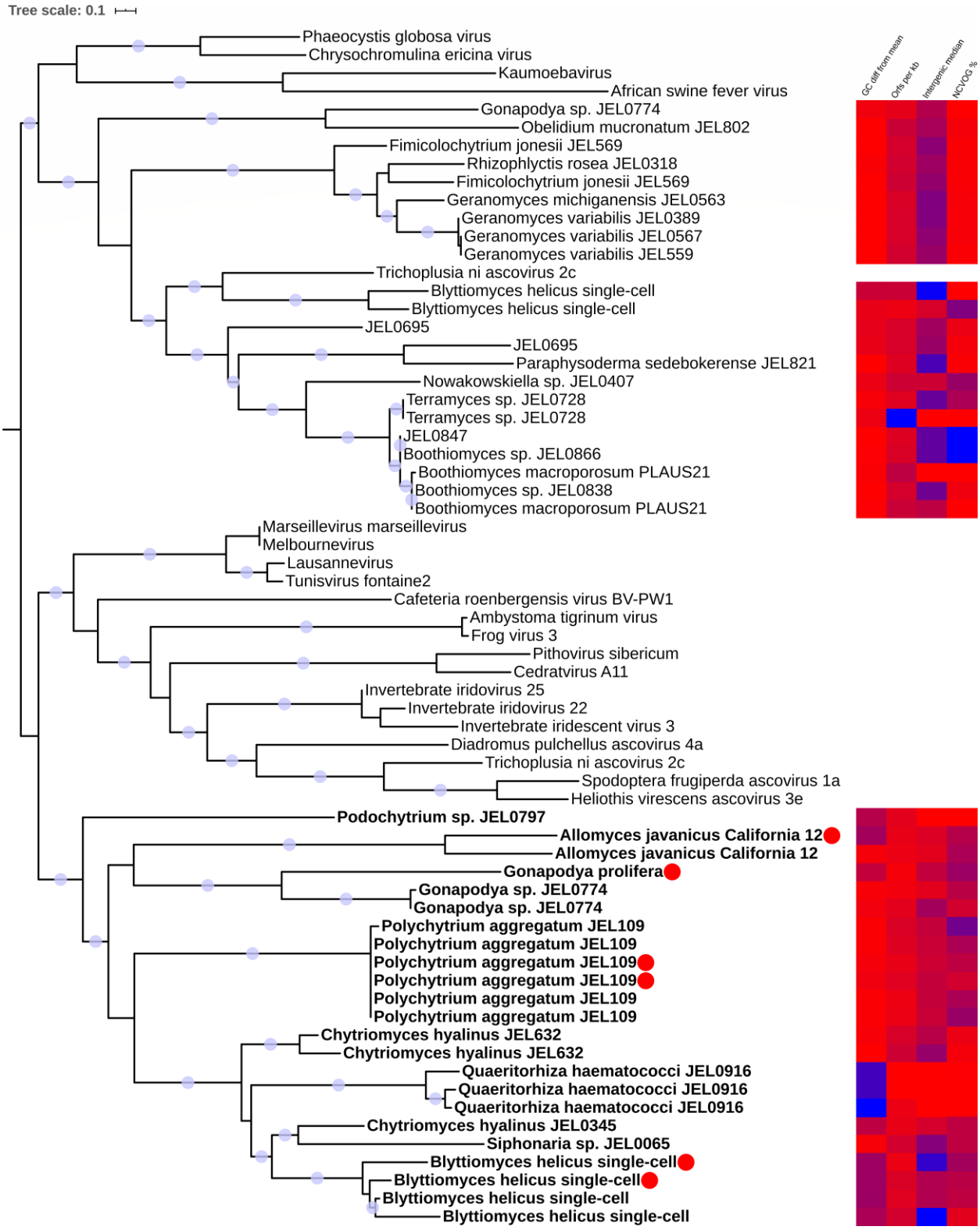
Supplementary Figure 4.6 Gene tree of NCVOG0211 (myristylated IMV envelope protein). Only one gene met our criteria, indicate by the boldened leaf. Nodes with ≥ 50% bootstrap support are indicated by a blue dot.



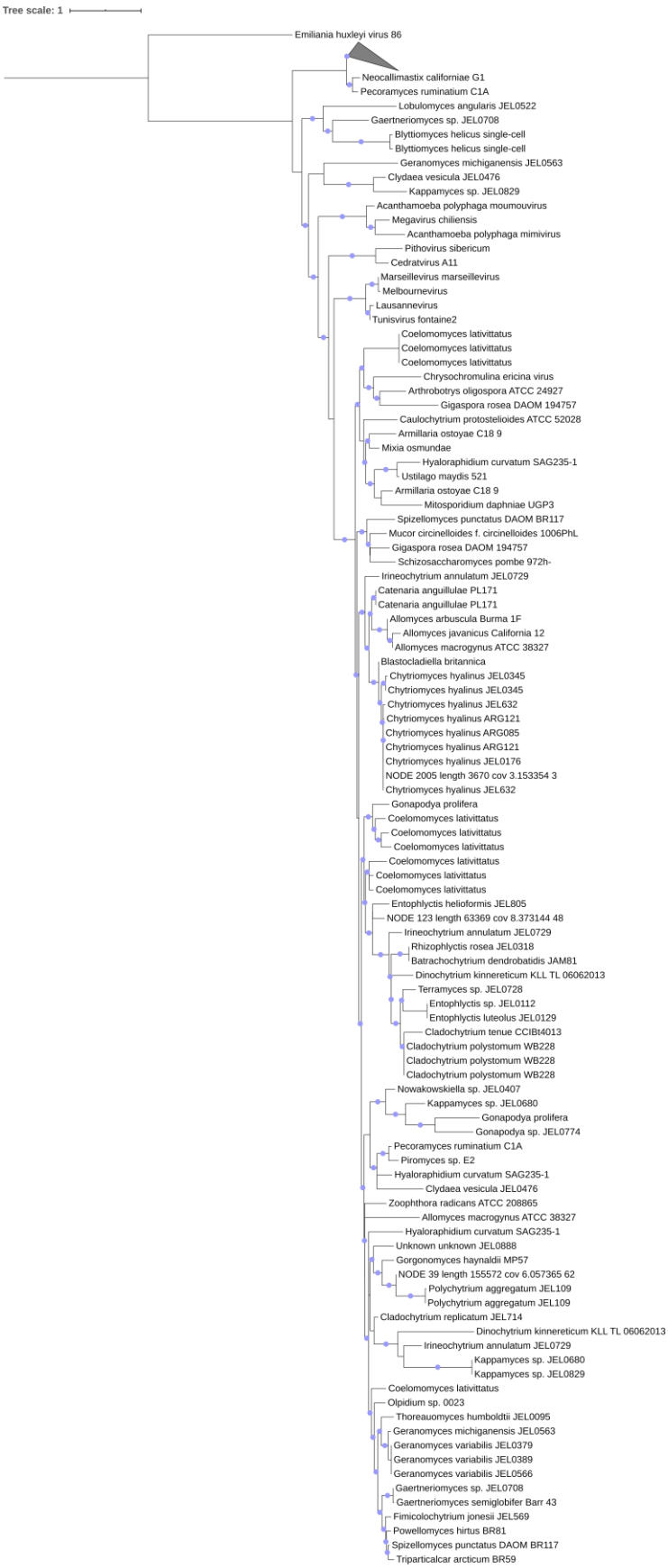
Supplementary Figure 4.7 Gene tree of NCVOG0249 (A32-like packaging ATPase). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, ORFs per kb sequence, intergenic median, and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).


Tree scale: 0.1
Cedratvirus A11
Megavirus chiliensis
Megavirus chiliensis
Acanthamoeba polyphaga mimivirus
Acanthamoeba polyphaga moumouvirus
Kaumoebavirus
African swine fever virus
Emiliania huxleyi virus 86
Mollivirus sibericum
Hyaloraphidium curvatum SAG235-1
Hyaloraphidium curvatum SAG235-1
Lausannevirus
Tunisvirus fontaine2
Marseillevirus marseillevirus
Melbournevirus
Ambystoma tigrinum virus
Frog virus 3
Invertebrate iridovirus 22
Invertebrate iridescent virus 3
Invertebrate iridovirus 25
Diadromus pulchellus ascovirus 4a
Trichoplusia ni ascovirus 2c
Spodoptera frugiperda ascovirus 1a
Heliothis virescens ascovirus 3e
Ostreococcus tauri virus 1
Heterosigma akashiwo virus 01
Paramecium bursaria Chlorella virus 1
Ectocarpus siliculosus virus 1
Feldmannia species virus
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus JEL0176
Polychytrium aggregatum JEL109
Quaeritorhiza haematococci JEL0916
Chrysochromulina ericina virus
Phaeocystis globosa virus
Phaeocystis globosa virus
Cafeteria roenbergensis virus BV-PW1
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Allomyces javanicus California 12
Allomyces arbuscula Burma1F
Allomyces javanicus California 12
Allomyces javanicus California 12
Allomyces javanicus California 12
Allomyces javanicus California 12
Gonapodya prolifera
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Neocallimastix californiae G1
Neocallimastix californiae G1
Anaeromyces robustus
Piromyces sp. E2
Piromyces finnis
GC diff from mean
Orfs per kb
Intergenic median
NCVOG %

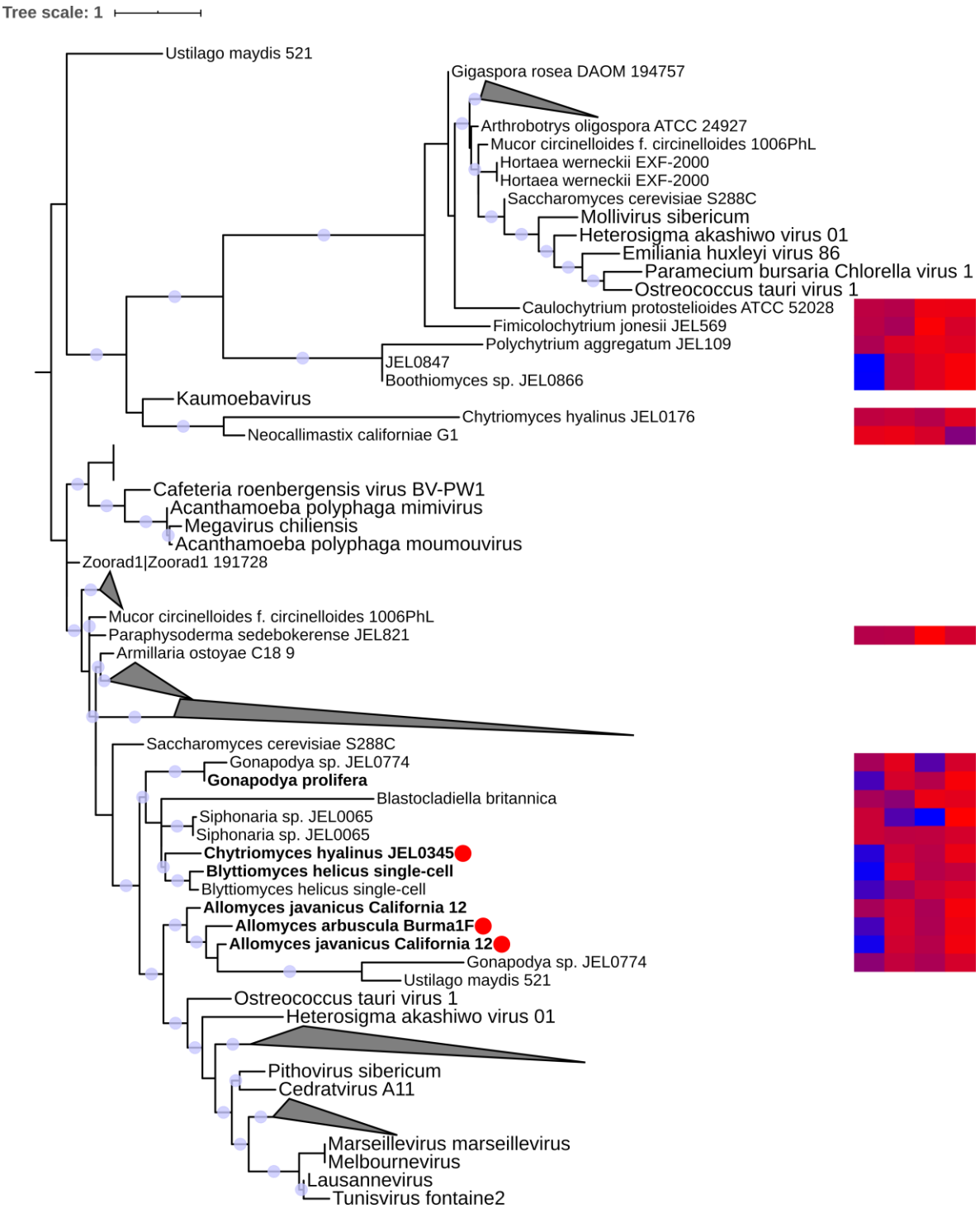
Supplementary Figure 4.8 Gene tree of NCVOG1088 (RNA ligase). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with $\geq$ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, ORFs per kb sequence, intergenic median, and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).


Tree scale: 0.1
GC diff from mean
Orfs per kb
Intergenic median
NCVOG %
Phaeocystis globosa virus
Chrysochromulina ericina virus
Kaumoebavirus
African swine fever virus
Gonapodya sp. JEL0774
Obelidium mucronatum JEL802
Fimicolochytrium jonesii JEL569
Rhizophlyctis rosea JEL0318
Fimicolochytrium jonesii JEL569
Geranomyces michiganensis JEL0563
Geranomyces variabilis JEL0389
Geranomyces variabilis JEL0567
Geranomyces variabilis JEL559
Trichoplusia ni ascovirus 2c
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
JEL0695
JEL0695
Paraphysoderma sedebokerense JEL821
Nowakowskiella sp. JEL0407
Terramyces sp. JEL0728
Terramyces sp. JEL0728
JEL0847
Boothiomyces sp. JEL0866
Boothiomyces macroporosum PLAUS21
Boothiomyces sp. JEL0838
Boothiomyces macroporosum PLAUS21
Marseillevirus marseillevirus
Melbournevirus
Lausannevirus
Tunisvirus fontaine2
Cafeteria roenbergensis virus BV-PW1
Ambystoma tigrinum virus
Frog virus 3
Pithovirus sibericum
Cedratvirus A11
Invertebrate iridovirus 25
Invertebrate iridovirus 22
Invertebrate iridescent virus 3
Diadromus pulchellus ascovirus 4a
Trichoplusia ni ascovirus 2c
Spodoptera frugiperda ascovirus 1a
Heliothis virescens ascovirus 3e
Podochytrium sp. JEL0797
Allomyces javanicus California 12
Allomyces javanicus California 12
Gonapodya prolifera
Gonapodya sp. JEL0774
Gonapodya sp. JEL0774
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Chytriomyces hyalinus JEL632
Chytriomyces hyalinus JEL632
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Quaeritorhiza haematococci JEL0916
Chytriomyces hyalinus JEL0345
Siphonaria sp. JEL0065
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell

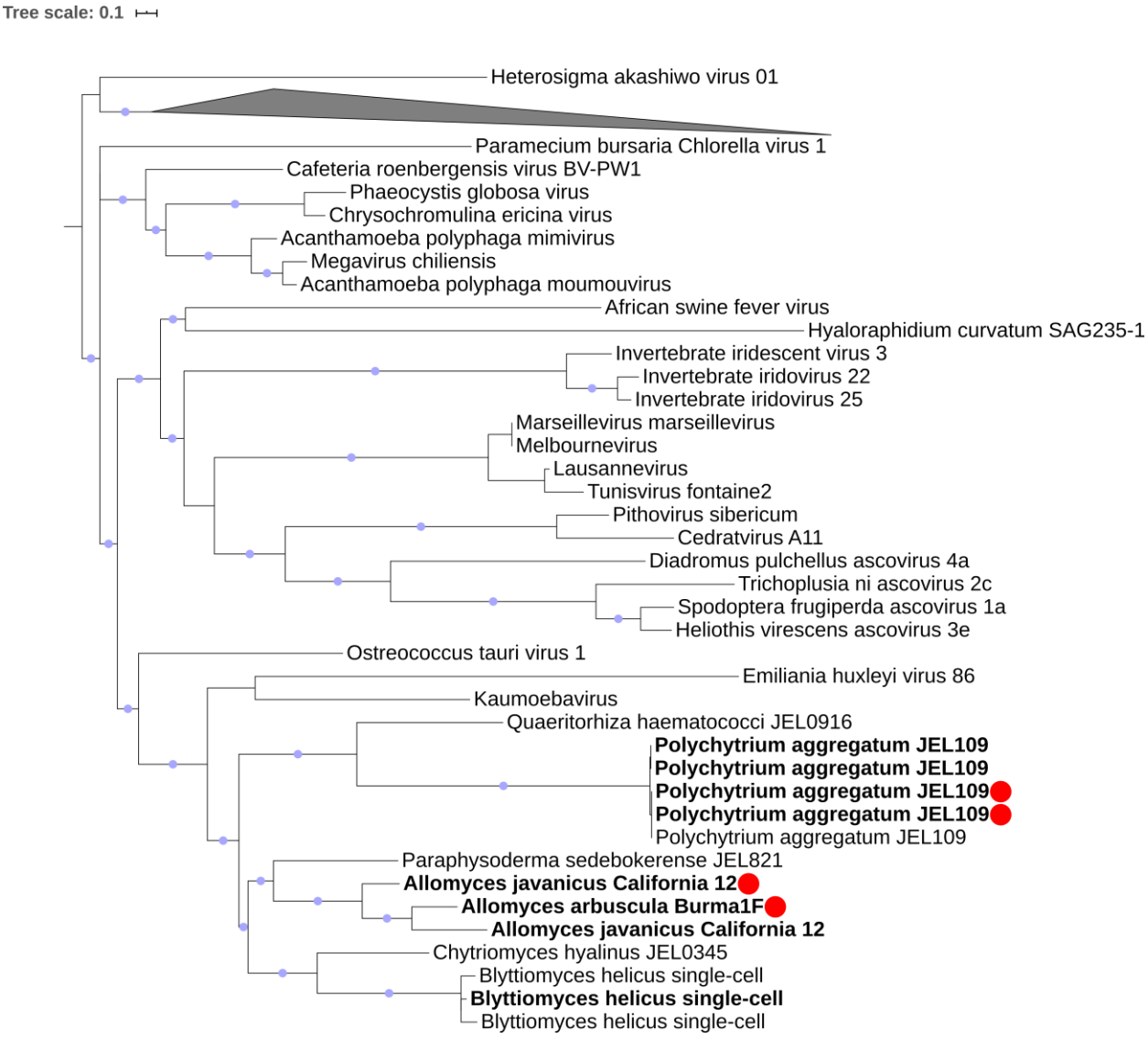
Supplementary Figure 4.9 Gene tree of NCVOG1115 (uracil-DNA glycosylase). No genes in our analyses met the criteria for consideration as fungal NCLDV genes, suggesting that this is not an NCVOG conserved in *Mycodnaviridae.* Nodes with ≥ 50% bootstrap support are indicated by a blue dot.


Tree scale: 1
Emiliania huxleyi virus 86
Neocallimastix californiae G1
Pecoramyces ruminatium C1A
Lobulomyces angularis JEL0522
Gaertneriomyces sp. JEL0708
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Geranomyces michiganensis JEL0563
Clydaea vesicula JEL0476
Kappamyces sp. JEL0829
Acanthamoeba polyphaga moumouvirus
Megavirus chiliensis
Acanthamoeba polyphaga mimivirus
Pithovirus sibericum
Cedratvirus A11
Marseillevirus marseillevirus
Melbournevirus
Lausannevirus
Tunisvirus fontaine2
Coelomomyces lativittatus
Coelomomyces lativittatus
Coelomomyces lativittatus
Chrysochromulina ericina virus
Arthrobotrys oligospora ATCC 24927
Gigaspora rosea DAOM 194757
Caulochytrium protostelioides ATCC 52028
Armillaria ostoyae C18 9
Mixia osmundae
Hyaloraphidium curvatum SAG235-1
Ustilago maydis 521
Armillaria ostoyae C18 9
Mitosporidium daphniae UGP3
Spizellomyces punctatus DAOM BR117
Mucor circinelloides f. circinelloides 1006PhL
Gigaspora rosea DAOM 194757
Schizosaccharomyces pombe 972h-
Irineochytrium annulatum JEL0729
Catenaria anguillulae PL171
Catenaria anguillulae PL171
Allomyces arbuscula Burma 1F
Allomyces javanicus California 12
Allomyces macrogynus ATCC 38327
Blastocladiella britannica
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus JEL632
Chytriomyces hyalinus ARG121
Chytriomyces hyalinus ARG085
Chytriomyces hyalinus ARG121
Chytriomyces hyalinus JEL0176
NODE 2005 length 3670 cov 3.153354 3
Chytriomyces hyalinus JEL632
Gonapodya prolifera
Coelomomyces lativittatus
Coelomomyces lativittatus
Coelomomyces lativittatus
Coelomomyces lativittatus
Coelomomyces lativittatus
Coelomomyces lativittatus
Entophlyctis helioformis JEL805
NODE 123 length 63369 cov 8.373144 48
Irineochytrium annulatum JEL0729
Rhizophlyctis rosea JEL0318
Batrachochytrium dendrobatidis JAM81
Dinochytrium kinnereticum KLL TL 06062013
Terramyces sp. JEL0728
Entophlyctis sp. JEL0112
Entophlyctis luteolus JEL0129
Cladochytrium tenue CCIBt4013
Cladochytrium polystomum WB228
Cladochytrium polystomum WB228
Cladochytrium polystomum WB228
Nowakowskiella sp. JEL0407
Kappamyces sp. JEL0680
Gonapodya prolifera
Gonapodya sp. JEL0774
Pecoramyces ruminatium C1A
Piromyces sp. E2
Hyaloraphidium curvatum SAG235-1
Clydaea vesicula JEL0476
Zoophthora radicans ATCC 208865
Allomyces macrogynus ATCC 38327
Hyaloraphidium curvatum SAG235-1
Unknown unknown JEL0888
Gorgonomyces haynaldii MP57
NODE 39 length 155572 cov 6.057365 62
Polychytrium aggregatum JEL109
Polychytrium aggregatum JEL109
Cladochytrium replicatum JEL714
Dinochytrium kinnereticum KLL TL 06062013
Irineochytrium annulatum JEL0729
Kappamyces sp. JEL0680
Kappamyces sp. JEL0829
Coelomomyces lativittatus
Olpidium sp. 0023
Thoreauomyces humboldtii JEL0095
Geranomyces michiganensis JEL0563
Geranomyces variabilis JEL0379
Geranomyces variabilis JEL0389
Geranomyces variabilis JEL0566
Gaertneriomyces sp. JEL0708
Gaertneriomyces semiglobifer Barr 43
Fimicolochytrium jonesii JEL569
Powellomyces hirtus BR81
Spizellomyces punctatus DAOM BR117
Triparticalcar arcticum BR59

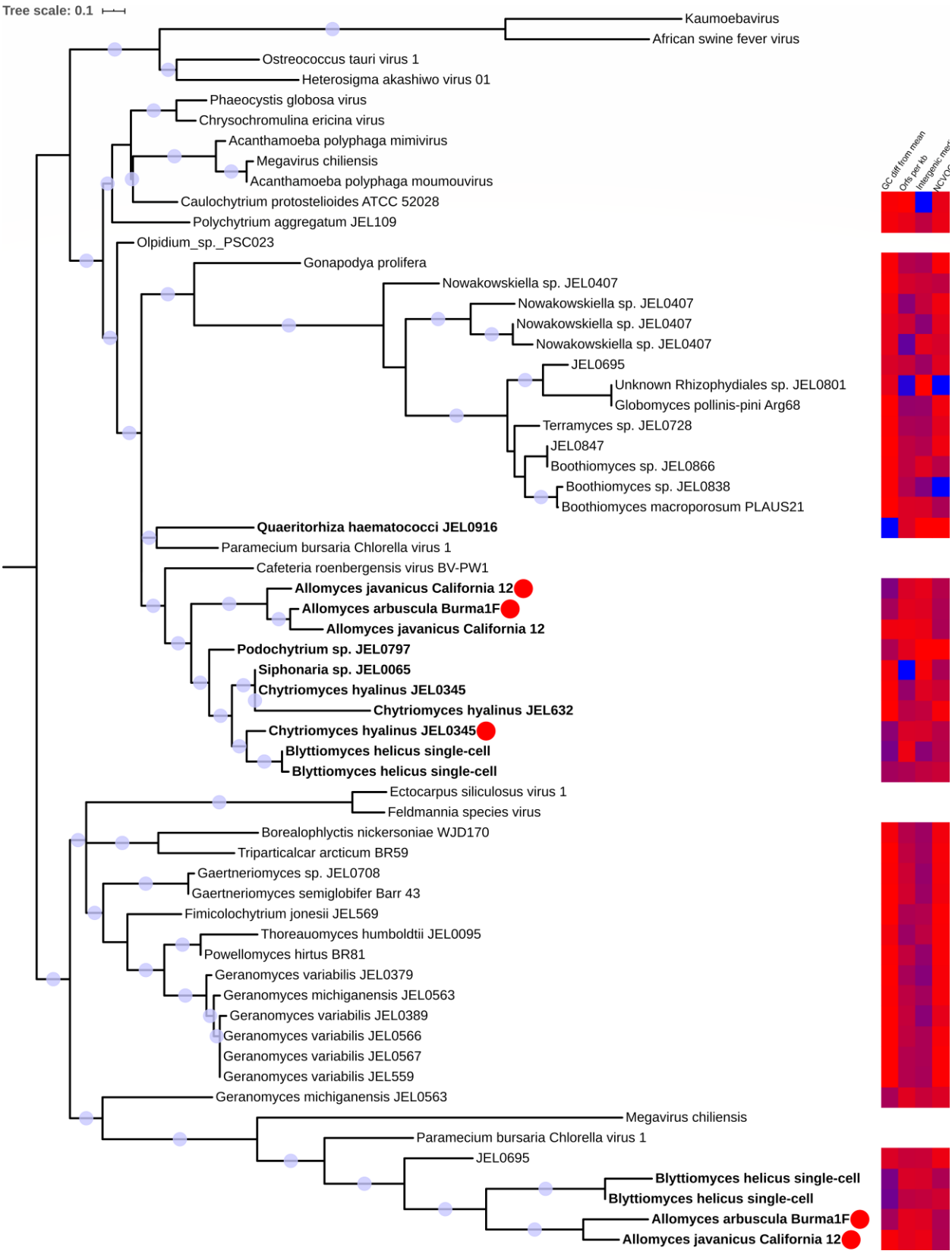
Supplementary Figure 4.10 Gene tree of NCVOG1117 (mRNA capping enzyme large subunit). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, ORFs per kb sequence, intergenic median, and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).



Supplementary Figure 4.11 Gene tree of NCVOG1127 (transcription initiation factor IIB). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot.



Supplementary Figure 4.12 Gene tree of NCVOG1192 (YqaJ viral recombinase family). Bolded leaves are fungal NCLDV genes that met our criteria. Red circles indicate contigs that also contain a copy of NCVOG0022 (major capsid protein). Nodes with ≥ 50% bootstrap support are indicated by a blue dot. Heatmap displays contig statistics: absolute value of %GC difference from mean, ORFs per kb sequence, intergenic median, and % contig hitting to NCVOG (evalue 1e-10) (left to right; blue = max value, red = min value).


Tree scale: 0.1
Kaumoebavirus
African swine fever virus
Ostreococcus tauri virus 1
Heterosigma akashiwo virus 01
Phaeocystis globosa virus
Chrysochromulina ericina virus
Acanthamoeba polyphaga mimivirus
Megavirus chiliensis
Acanthamoeba polyphaga moumouvirus
Caulochytrium protostelioides ATCC 52028
Polychytrium aggregatum JEL109
Olpidium_sp._PSC023
Gonapodya prolifera
Nowakowskiella sp. JEL0407
Nowakowskiella sp. JEL0407
Nowakowskiella sp. JEL0407
Nowakowskiella sp. JEL0407
JEL0695
Unknown Rhizophydiales sp. JEL0801
Globomyces pollinis-pini Arg68
Terramyces sp. JEL0728
JEL0847
Boothiomyces sp. JEL0866
Boothiomyces sp. JEL0838
Boothiomyces macroporosum PLAUS21
Quaeritorhiza haematococci JEL0916
Paramecium bursaria Chlorella virus 1
Cafeteria roenbergensis virus BV-PW1
Allomyces javanicus California 12
Allomyces arbuscula Burma1F
Allomyces javanicus California 12
Podochytrium sp. JEL0797
Siphonaria sp. JEL0065
Chytriomyces hyalinus JEL0345
Chytriomyces hyalinus JEL632
Chytriomyces hyalinus JEL0345
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Ectocarpus siliculosus virus 1
Feldmannia species virus
Borealophlyctis nickersoniae WJD170
Triparticalcar arcticum BR59
Gaertneriomyces sp. JEL0708
Gaertneriomyces semiglobifer Barr 43
Fimicolochytrium jonesii JEL569
Thoreauomyces humboldtii JEL0095
Powellomyces hirtus BR81
Geranomyces variabilis JEL0379
Geranomyces michiganensis JEL0563
Geranomyces variabilis JEL0389
Geranomyces variabilis JEL0566
Geranomyces variabilis JEL0567
Geranomyces variabilis JEL559
Geranomyces michiganensis JEL0563
Megavirus chiliensis
Paramecium bursaria Chlorella virus 1
JEL0695
Blyttiomyces helicus single-cell
Blyttiomyces helicus single-cell
Allomyces arbuscula Burma1F
Allomyces javanicus California 12
GC diff from mean
Orfs per kb
Intergenic mean

Conclusion

The early-diverging lineages of fungi have challenged and transformed our knowledge of the fungal kingdom. Similarly, plant and fungal viruses have challenged and transformed conventional understandings of the roles of viruses. Mycoviruses in early-diverging fungi have, indeed, challenged what we know about fungal viruses. Prior to my dissertation research, the intersection of these two fields was untraversed territory whose exploration would fill in the map in unexpected ways.

In the course of this dissertation, I have argued that viruses may be more prevalent in the basal fungi relative to the Dikarya, which could implicate viruses as forces in the evolution of fungal traits such as septa and life histories devoid of zoosporic stages. I found that fungi which had been maintained in culture for decades were harboring mycoviruses, invisible to the researchers who preserved them. Many of these cultures belong to research collections that are disseminated to laboratories worldwide, for multifarious research purposes, unbeknownst of their infection status.

Such mycoviral persistence and, often, lack of phenotype resulting from infection has prompted hypotheses of ancient coevolution of mycoviruses and their hosts. Due to a lack of diverse fungal sampling, tests of evolutionary relationships could not previously be addressed. My statistical analyses, which included new mycoviral sequence data from hosts that spans the breadth of the fungal kingdom, demonstrate significant evidence for cospeciation of mycoviruses

and fungi. Importantly, cospeciation is rare and is perhaps suggestive of unique viral-host interactions.

Known mycoviruses are almost exclusively composed of RNA genomes. Perhaps most exciting, my work has more than quadrupled the number of known viruses with DNA genomes in fungi with the discovery of nucleocytoplasmic large DNA viruses (NCLDV) in basal fungal lineages. This finding has unknown, but possibly important, implications for global nutrient cycling, as large DNA viruses notoriously reprogram their host's metabolism. This finding opens new avenues for research which be fruitful for years to come. Metabolomics, particularly, is an exciting frontier which could bring to the surface some of these important viral effects on the host, which have ecosystem-level implications. There exists untapped and previously unappreciated potential for new model systems to be developed to probe virus-host interactions across taxonomic and ecological scales. In developing these model systems, more scientists may be "lured by the siren's song" in the early-diverging lineages, and thus help ensure the study of these organims for generations to come.

Marine virologists have been in on a secret, articulated by Brussaard and colleagues in 2008, "...the effect viruses have on organisms… makes them the ultimate nanoscale drivers/regulators of life." Perhaps unironically, this dissertation has been completed during one of the most significant pandemics in modern history, with the ultimate culprit a virus. If it wasn't clear before, it is viscerally obvious now that, indeed, viruses are the ultimate regulators of life. As we consider the intricate and integrated ways that viruses regulate life, at the individual, population, and ecosystem scales, we must actively ensure that a diversity of organisms are included in our work. Perhaps more than anything else, this dissertation is a reminder that representation matters.