Global distribution and richness of *Armillaria* species inferred from public databases and amplicon sequencing datasets

Rachel A. Koch[1] and Joshua R. Herr[1,2,*]

[1]Department of Plant Pathology, University of Nebraska, Lincoln, NE 68503, USA; [2]Center for Plant Science Innovation, University of Nebraska, Lincoln, NE 68588, USA; [*]Correspondence: Joshua R. Herr (jherr@unl.edu)

ABSTRACT

Armillaria is a globally distributed fungal genus most notably recognized as economically important plant pathogens that are found predominantly in forest and agronomic systems. *Armillaria* has more recently received attention for ecologically diverse roles as woody plant decomposers and as mycorrhizal symbionts with specific plants. Previous phylogenetic analyses suggest that around 50 species are recognized globally. Despite this previous work, no studies have analyzed the global species richness and distribution of the genus using data derived from fungal community sequencing datasets. In order to assess the global diversity and species richness of *Armillaria*, we mined publicly available sequencing datasets derived from numerous primer pairs for the ribosomal operon. Our estimates reveal that species richness ranges from 50 to 60 species, depending on which marker is used. Eastern Asia represents the biogeographic region with the highest species richness. We also assess the overlap of species across geographic regions and propose some hypotheses regarding the drivers of variability in species diversity and richness between different biogeographic regions.

Keywords: DNA Barcoding; Basidiomycete fungi; Diversification; Meta-analysis; Plant pathogenic fungi; Ribosomal operon; Symbiosis

INTRODUCTION

While the importance that fungi play as mutualists, decomposers, and pathogens is undisputed, scientists are just beginning to disentangle the processes that shape their global species richness and distribution (Tedersoo et al., 2014). One such globally important group is the armillarioid clade—a lineage composed of the mushroom-forming genera *Armillaria* and *Desarmillaria,* along with the sequestrate genus *Guyanagaster*. Broadly speaking, the relative species richness and distribution patterns of these genera are understood. Both of the two described species of *Guyanagaster* have a known distribution limited to the Guiana Shield of northern South America (Henkel et al., 2010; Koch et al., 2017). *Desarmillaria* species, representative of the exannulate mushroom-forming species within the armillarioid clade, also contains two described species, one of which is restricted to northern European bogs (Zolciak et al., 1997), and the other which is pathogenic and presumably widely distributed throughout the northern hemisphere (Guillaumin et al., 1993). In contrast, *Armillaria* species, representative of the annulate mushroom-forming species within the armillarioid clade, are globally distributed with an estimated 40 to 50 species (Volk & Burdsall, 1995; Baumgartner et al., 2011). While *Guyanagaster* is a member of the armillarioid clade (Koch et al., 2017), herein the discussion of the armillarioid clade will only refer to species of *Armillaria* and *Desarmillaria*, whose ecology and pathology are better studied, and when necessary, *Guyanagaster* will be specifically referenced.

Armillarioid species are recognized as ecologically and economically important plant pathogens, identified as the causal agents of Armillaria Root Rot. As facultative necrotrophs, these species possess both parasitic and saprotrophic phases. During their parasitic phase, species may attack a broad range of woody dicotyledonous hosts in both natural and agronomic systems (reviewed in Baumgartner et al., 2011). However, during their saprotrophic phase, armillarioid species play an important role in carbon cycling and nutrient exchange as they are able to degrade the entire spectrum of the major components of plant cell walls (Kile et al., 1991; Sipos et al., 2017). One largely diagnostic characteristic of many armillarioid species is the production of rhizomorphs—discrete, filamentous structures that forage for new resources away from the primary resource base (Garraway et al., 1991). Rhizomorphs can grow long distances to infect plants, resulting in genets with exceptionally large sizes, some of which have been recognized as the largest and oldest organisms on earth (Smith et al., 1992).

In addition to their importance as plant pathogens and biomass decomposers, armillarioid species also have the genetic capacity to engage in several symbiotic associations with various plants, fungi, and insects. For example, the growth of sclerotia of the polyporoid fungus, *Polyporus umbellatus*, is dependent on nutrition from symbiotic *Armillaria* species (Guo & Xu, 1992). *Armillaria* species in east Asia engage in a form of mycorrhizal relationship known as mycoheterotrophy with species of achlorophyllus orchids in the genera *Galeola* and *Gastrodia* (Kikuchi & Yamaji, 2010; Guo et al., 2016). In other examples, *Armillaria* sporocarps are parasitized by the mushroom-forming species *Entoloma abortivum* (Lindner Czederpiltz et al., 2001; Koch & Herr, 2021) and potentially also by species in the Ascomycete genus *Wynnea* (Xu et al., 2019). Lastly, *Guyanagaster necrorhizus* harbors nitrogen-fixing bacterial species within

its sporocarps, which likely facilitates a dispersal mutualism with wood-feeding termites (Koch et al., 2018; Koch et al., 2021).

Traditionally, species within the armillarioid clade were delimited based on morphological characters and mating compatibility studies (Bérubé & Dessureault, 1989; Cha et al., 1994; Volk & Burdsall, 1995; Qin et al., 2007). However, given the difficulty in identifying species based solely on these features, there has been significant emphasis on understanding species boundaries within the armillarioid clade using phylogenetic data (Guo et al., 2016; Klopfenstein et al., 2017; Koch et al., 2017; Coetzee et al., 2018). Early studies relied largely on data from the nuclear ribosomal cistron — which bridges the small subunit (18S), internal transcribed spacer region 1 (ITS1), 5.8S subunit, internal transcribed spacer region 2 (ITS2), the intergenic spacer region (IGS), and the large subunit (28S). However, inadequate resolution of these markers led to the hunt for markers with more accurate phylogenetic resolution. The *translation elongation factor 1* (*tef1*) represents the most widely adopted marker, in that it provides good resolution to infer species boundaries and lineage relationships within a global context (Maphosa et al., 2006; Guo et al., 2016; Klopfenstein et al., 2017; Coetzee et al., 2018).

While *tef1* is a good marker for species discrimination in the armillarioid clade, it is not widely utilized in fungal barcoding initiatives or community sequencing studies. Instead, markers that have consensus from the fungal research community for such studies include the ITS1 and ITS2 region (Schoch et al., 2011; Herr et al., 2013; Hibbett et al., 2015). Because the variable regions of the ITS are so widely used in the aforementioned types of studies, we wanted to assess—from the perspective of these markers—the overall global distribution and species richness within the

armillarioid clade, much like has been done for fungi in general (Tedersoo et al., 2014). To do so, we analyzed the totality of publicly available ITS sequences submitted to NCBI-GenBank and fungal community sequencing data submitted to NCBI-SRA to gain a more complete understanding of the species richness and the global distribution of the armillarioid clade. This also allowed us to utilize data generated during the course of fungal barcoding initiatives (Truong et al., 2017; Haelewaters et al., 2018; Vu et al., 2019). Additionally, we speculate as to how symbiotic interactions of armillarioid species affect the diversity and distribution of this lineage.

MATERIALS AND METHODS

Sequences curated from public database vouchers. In order to create a dataset of all publicly available armillarioid ITS sequences, we queried NCBI-GenBank (https://www.ncbi.nlm.nih.gov) for "*Armillaria*" and "ITS" (or "*Desarmillaria*" and "*Guyanagaster*" with "ITS") and downloaded all of the nucleotide sequences that matched these criteria. Taxonomy Browser (NCBI) (Federhen, 2012) was then used to verify the correct taxonomic placement for each of these sequences. The feature for geographic region (such as "country") was also extracted for each sequence. Any sequences without this corresponding metadatum were excluded. Finally, all sequences with the identifier "*Armillaria*" from the UNITE v8.3 database (Nilsson et al., 2018) were also included in the dataset irrespective if they had corresponding location metadata. For all of these sequences, we then extracted the ITS1 and ITS2 regions using the ITSx tool (Bengtsson-Palme et al., 2013).

Sequences curated from public amplicon studies. Next, we mined the Sequence Read Archive (SRA) database for public datasets of fungal amplicon sequencing data of the ITS1 and ITS2 region. We identified 454 pyrosequencing and various Illumina platform sequencing datasets with the query (ITS2, ITS1 AND "biomol dna"[Properties] AND "platform ls454"[Properties] AND “platform Illumina”[Properties]) to the SRA database (Leinonen et al., 2010) which provided us with 183,838 sample files representing 1,253 public deposits to the SRA. We then manually curated a list of datasets based on the availability of metadata from the NCBI SRA Run Selector, such as geographic location, as we did previously for the NCBI sequences. Once datasets were downloaded in the “.sra” format using the `prefetch` command, we converted the datasets to “.fastq” format using the `fasterq-dump` command with the “--split_files” option. Both tools are available from NCBI through the sra-tools scripts (https://github.com/ncbi/sra-tools). Sequences from SRA deposited datasets were clustered against the NCBI list for *Armillaria* (generated in the first stage) using VSEARCH version 2.17.0 (Rognes et al., 2016) at 90% sequence similarity using the `--usearch_global` command. In order to avoid off-target clustering and to double check the VSEARCH clustering quality, we next identified the output from the VSEARCH clustering using DECIPHER version 3.13 IDTAXA algorithm `Classify Organisms` (Wright, 2016) and the UNITE2020 DECIPHER database (http://www2.decipher.codes/Classification/TrainingSets/UNITE_v2020_February2020.RData) selecting only those sequences being taxonomically identified as *Armillaria*. Out of the 1,253 public studies deposited to the SRA that we queried, only 56 studies, restricted to forest soil and plant matter, contained sequences with 90% homologous base pairs to *Armillaria* ITS sequences. The remaining 1,197 studies, which mainly consisted of ecosystems that did not contain woody

plant hosts (i.e. built-environment, human, etc.), did not contain any sequences with homology to existing *Armillaria* sequences.

Curation of total sequence datasets and geographic metadata. Geographic location metadata were also extracted from studies containing *Armillaria* sequences, which represented *Armillaria* voucher specimens deposited in NCBI Genbank and from the SRA database metadata files. Sequences of the ITS1 region from the Genbank nucleotide database and the SRA were combined, and the same was done as a separate dataset for the ITS2 region. Any sequences less than 150 base pairs in length were removed. To estimate species from these sequences, we then clustered the ITS1 and ITS2 regions (combining the voucher specimen data and SRA sequences) at 98% using USEARCH 11 (Edgar, 2010). We chose 98% nucleotide identity threshold because this was the identity threshold utilized by Tedersoo et al. (2014). The biogeographic regions we refer to here largely follow those utilized by Tedersoo et al. (2014). Because some of the armillarioid sequences associated with *Wynnea* species did not pass our quality filters, and we were still interested in utilizing them, we independently blasted each of them to our dataset to place them in the appropriate species cluster.

Phylogenetic analysis. To understand and visualize how the armillarioid species clusters were related we utilized a phylogenetic approach. One exemplar sequence from each species cluster was chosen to include in the multiple sequence alignment. We utilized the full length sequence and not just the hypervariable regions, when possible. Species clusters for the ITS1 region and ITS2 region were analyzed separately. These sequences were aligned using MAFFT version 7 (Katoh & Standley 2013) with refinements to the alignment performed manually. RAxML-NG

(Kozlov et al. 2019) was used to reconstruct this phylogeny. The two known species of *Guyanagaster*, sister to *Armillaria* and *Desarmillaria* species (Koch et al. 2017), were used to root this phylogeny. Well supported branches are defined as having 70% or greater bootstrap support. The location metadata associated with each species cluster was mapped onto each phylogenetic tree using the `ggtree` package (Yu et al. 2018) within the R 4.1.0 environment (R Core Team 2013).

RESULTS

A total of 1,153 *Armillaria* ITS1 region sequences (1,119 from the nucleotide database and 34 from the SRA database) and 1,567 *Armillaria* ITS2 region sequences (1,246 from the nucleotide database and 321 from the SRA database) were sorted into species clusters. Public data sets which contributed to the high-throughput ITS1 and ITS2 region sequences in this study consisted of six different primer pair combinations: fITS7/ITS4, gITS7/ITS4, ITS1F/ITS2, ITS1F/ITS4, ITS3/ITS4, and ITS5/ITS2. Sequences for both ITS marker regions were identified from samples collected from all continents except Antarctica. At a 98% nucleotide identity, 51 armillaroid species clusters were recovered for the ITS1 region and 61 for ITS2 region. The distribution of these species clusters are shown in Fig. 1. Of the ITS1 region species clusters, sequences from the SRA database clustered into five species clusters (species clusters 5, 6, 9, 15 and 23), all of which also contained sequences queried from the NCBI nucleotide database. Of the ITS2 species clusters, sequences derived from the SRA database were clustered into 14 species at 98% similarity (species clusters 3, 9, 10, 13, 17, 18, 19, 23, 24, 26, 30, 35, 43 and 51); one of these clusters, species cluster 17, was solely composed of sequences from the SRA database and was not clustered with any sequences from the NCBI nucleotide database.

The most species rich biogeographic region was East Asia, with 23 and 19 species clusters at ITS1 and ITS2, respectively (Fig. 1 and 2). With the ITS1 region data, East Asia also contained the most species clusters that are found in no other biogeographic region (Fig. 2). The other species rich biogeographic regions included western and eastern North America and Europe, with five, six, and nine armillarioid species clusters at the ITS1 region and 13, 12, and 12 armillarioid species clusters at the ITS2 region, respectively. However, it is clear that a majority of the total barcoded armillarioid sporocarps, rhizomorphs, and mycelial tissue are from these regions (Fig. 3). In the southern hemisphere, species richness was less compared to the northern hemisphere, with temperate South America, Africa, Australia, and New Zealand having five, six, three, and four species clusters at the ITS1 region and eight, five, six, and nine species clusters at the ITS2 region. Largely tropical geographic areas, including Hawaii, India, tropical South America and Southeast Asia had the lowest species richness with one to two mushroom-forming species in each location. Tropical South America also has two species of the sequestrate genus, *Guyanagaster*.

Shared species clusters between biogeographic regions are much more pervasive in the northern hemisphere compared to the southern hemisphere. Between locations, the most shared species clusters include a link between East Asia and Europe, as well as a link between Australia and New Zealand. Evidence for shared species clusters throughout the southern hemisphere is limited to only Australia and New Zealand which contain several identical species clusters (Fig. 2), but do not share any species with any other geographic location (Fig. 1). At the ITS1 region, four species clusters (species clusters 1, 4, 13, and 22), and at the ITS2 region, three species (species

clusters 1, 7, and 12), were shared between locations in the northern and southern hemisphere. Two of those species clusters derived from the ITS1 region and all of the species clusters from the ITS2 region included sequences from specimens or environmental samples in the northern hemisphere and Subsaharan Africa.

Species clusters (or species complexes) for the armillarioid lineage were determined from the combined dataset of sequences derived from community sequencing efforts and accessions deposited into NCBI Genbank. We detected the two known sequestrate armillarioid species—species of *Guyanagaster*—from the nucleotide database. Of the *Desarmillaria* species, there was one species cluster detected at both the ITS1 (species cluster 14) and ITS2 (species cluster 22) regions, from Europe and East Asia, which likely corresponds to *D. ectypa*. Four and three species clusters that correspond to *D. tabescens* at the ITS1 (species clusters 5, 6, 11 20) and ITS2 (species clusters 9, 25 and 41) regions, respectively, were also detected. These *Desarmillaria* species clusters form a well-supported monophyletic lineage in both analyses. All of the other species clusters, 45 at the ITS1 region and 56 at the ITS2 region, likely correspond to *Armillaria sensu stricto*. There were 11 species clusters at the ITS1 (species clusters 4, 7, 9, 19, 23, 26, 28, 29, 33, 39 and 41) region and ten at the ITS2 (species clusters 7, 12, 16, 19, 27, 48, 49, 52, 54 and 55) region that included sequences annotated as *A. mellea*. These *Armillaria mellea* species clusters, along with some *Wynnea* symbionts, form a well-supported monophyletic lineage in both analyses (Fig. 3), indicative that this is a species complex illustrated in the phylogenetic trees in Fig. 3. Species clusters for both the ITS1 and ITS2 regions are available in Tables S1 and S2.

DISCUSSION

Previous studies within the armillarioid clade have used morphology (Bérubé & Dessureault, 1988; Watling et al., 1991; Volk & Burdsall, 1995), hyphal compatibility tests (Anderson & Ullrich, 1979; Banik & Burdsall, 1998) and phylogenetic assessment (Maphosa et al., 2006; Guo et al., 2016; Coetzee et al., 2018) to draw conclusions of diversity and global richness. In this study, we utilized a complementary methodology — data derived from both the ITS1 and ITS2 marker regions taken from barcoding and community sequencing analysis studies — to investigate global species richness and distribution of species. One strength of this approach is that it allowed us to utilize the increasing numbers of publically available datasets (notably using the ITS1 and ITS2 marker regions) from fungal ribosomal operon barcoding efforts that have previously gone unused in *Armillaria* phylogenetic studies and diversity estimates. Previous studies addressing *Armillaria* phylogenetics have largely consisted of protein-coding sequences (Coetzee et al., 2011; Guo et al., 2016; Koch et al., 2017; Coetzee et al., 2018). Given our filtering scheme for short and spuriously aligned sequences, and the fact that we did not include any singletons -- recognized here as species clusters here that only contained a single sequence -- identified in the analysis, our results likely undercount the species richness of this clade. Additionally, as there are more environmental sequencing studies across the globe, particularly in undersampled regions, we will likely uncover greater species richness in this clade.

There are known biases with ITS as a marker within the armillarioid clade, including that it may be too conserved to discriminate some species (Chillali et al., 1998; Coetzee et al., 2001), or too variable and lead to an overestimation of individuals in other species complexes (Dunne et al., 2002). Although, when the hypervariable regions (ITS1 and ITS2) are considered, phylogentic

analysis for *Armillaria* performs relatively well (Chillali et al., 1998). As the universal barcode for fungi (Schoch et al., 2013) the ITS regions are the most widely available in terms of size and species representation in public databases (Nilsson et al., 2018) and the region is the predominant marker for fungal molecular ecology (Herr et al., 2013, Hibbett et al., 2016). Additionally, there has been an increasing awareness of the importance of barcoding macroscopic fungi, most notably those stored in fungal herbaria, using ITS primers to aid in identification and create more complete databases. However, given that other marker regions may be better at resolving patterns of armillarioid species richness, this study underscores the need for additional ecological markers for fungi and for the development of sequence analysis from fungal metagenomic and metatranscriptomic studies (Kuske et al., 2015; Hibbett et al., 2016). Because the aim of this work is to understand diversity and distribution patterns, and not to resolve nomenclatural issues within this clade, at present we do not know how our species clusters correspond with the species concepts derived from phylogenetic assessment. However, we do know that our findings here are largely consistent with findings utilizing phylogenetic analysis. In future studies, when using a greater number of markers, or a more phylogenetically informative marker region, we may also find that overall species richness of the armillarioid clade is greater than we report here. Lastly, it is important to note that the focus of this work is not meant to resolve any of the taxonomic or nomenclatural issues within this clade. We argue that this should be accomplished via a combination of morphological, ecological, and phylogenomic data.

Patterns of distribution and species richness

From the perspective of the ITS1 marker region, all of the armillarioid sequences derived from our curated SRA sequences clustered with sequences from the nucleotide database based on

voucher collected material. The ITS2 marker region is utilized more frequently in fungal community sequencing efforts due to its level of variability (Hibbett et al., 2016), so it is not surprising that there were more armillarioid species clusters detected from this marker region. The sequences from our curated amplicon sequencing dataset that are not represented by a voucher-derived sequence reads represent a single armillarioid species cluster from temperate South America, a geographic area warranting increased surveys for sporocarp diversity. Because several *Armillaria* species from that region do not have sequenced type or voucher specimens, it remains unknown if our armillarioid species cluster represents an undescribed species, or represents a species that has been described but has not been previously sequenced. Another interesting example that elucidates the importance of barcoding studies, includes species cluster 5 from the ITS1 region, as this cluster is composed of one sequence from a voucher specimen (Vu et al., 2019), with all the other sequences being from an environmental study with no associated vouchers. In some cases, it is difficult to discern which armillarioid species deposited to NCBI GenBank correspond to which species clusters from the total dataset and this appears to be largely due to misapplication of names deposited into GenBank (Bruns et al 2008). This further illuminates the importance of continued efforts to barcode macroscopic collections with well documented voucher material.

As an important pathogen in forest and agronomic systems there has been an emphasis on providing data to advance species delimitation of *Armillaria*. Additionally, the relatively characteristic signs of *Armillaria* presence in the environment, including white mycelial fans under tree bark and black shoestring-like rhizomorphs (Baumgartner et al., 2011) are easy to identify, collect, culture, and sequence species of *Armillaria*, which certainly aids in our

understanding of the distribution of these species even in the absence of sporocarps. Lastly, the entirety of the armillarioid clade has relatively few species when compared to other genera found in the Agaricales (Liimatainen et al., 2014; Oliveira et al., 2020). Of course, the available community sequencing studies from which we mined do not comprehensively represent all natural communities in which armillarioid species occur, so there are undoubtedly more species that are left to be found as sampling increases. However, these results also suggest that *Armillaria* species are not contributing substantially to the pool of species known based purely on sequence data (Hawksworth and Lücking, 2017) .

Our analyses here largely agree with previously phylogenetic assessments on diversity of this clade and which geographic regions harbor the greatest species richness in the armillarioid clade. This study shows that armillarioid species richness is concentrated in the northern hemisphere, specifically in East Asia. While this finding can certainly be attributed in part to an outsized collecting effort of this group throughout this geographic area (Fig. 3), it may also be reflective of the life history of this lineage. Eurasia is hypothesized to be the ancestral range of *Armillaria* (Koch et al., 2017) and these results are consistent with this region also being the center of species diversity. East Asia, and China in particular, was previously known to have a high species richness of *Armillaria* species (Qin et al., 2007; Zhao et al., 2008; Guo et al., 2016). Furthermore, phylogenetic analyses have provided significant evidence of genetic divergence between *Armillaria* specimens collected from China that were thought to be conspecific throughout the northern hemisphere, suggesting that endemicity may be prevalent in the northern hemisphere in this clade (Guo et al., 2016), and which our analyses also suggest to be true (Fig. 2).

Another feature of northern hemisphere diversity of the armillarioid clade is the high degree of connectedness between the three primary sampling regions (Fig. 1). Certain *Armillaria* species clusters in the northern hemisphere appear to be broadly distributed throughout the entire region. There are two and five *Armillaria* species clusters, respectively, that are present in North America, Europe and East Asia when considering the ITS1 and ITS2 markers. This finding supports the notion that some armillarioid species are distributed throughout the entire northern hemisphere (Coetzee et al., 2015; Coetzee et al., 2018). However, within the northern hemisphere, there are also many species that appear to be restricted to just a certain biogeographic region, suggesting that a narrower interpretation of specific species is more appropriate (Fig. 2). For example, the names *D. tabescens* and *A. mellea* have been applied to specimens collected throughout the northern hemisphere. However, our results suggest that these are species complexes (Fig. 3). This confirms a previous study that detected significant phylogenetic structure in specimens that match the description of *A. mellea* (Coetzee et al., 2000). While the same has been suspected for *D. tabescens*, inconclusive results from mating studies (Darmono et al., 1992; Guillaumin et al., 1993; Ota et al., 1998; Qin et al., 2007) and phylogenetic analyses (Tsykun et al., 2013; Coetzee et al., 2015), have clouded efforts to disentangle this complex.

Phylogenetic analyses exhibit strong support for an Australasian and temperate South American lineage (Coetzee et al., 2010; Koch et al., 2017; Coetzee et al., 2018). Species of *Armillaria* in both regions associate with species of *Nothofagus* (Coetzee et al., 2003a) and a dispersal passageway through Antarctica likely facilitated the observed evolutionary relationships of

species in this clade (Koch et al., 2017). Previous studies have suggested that some species, including *A. novae-zelandiae*, are shared between the two regions (Coetzee et al., 2018). However, results from our community analysis of these same data suggest that no armillarioid species are shared between temperate South America and Australasia (Fig. 1), and that morphologically similar species may be present in both geographic regions. This finding is supported by previous phylogenetic analyses of *Armillaria* species from these regions which shows considerable intraclade substructure from sporocarps in the different biogeographic regions (Pildain et al., 2009) and further suggests that dispersal between the two regions is not common. This could be due to the closure of overland dispersal routes throughout Antarctica several million years ago, and, as a result, the geographic barrier allowed for two lineages of armillarioid fungi to begin diversification. Finally, our analysis suggests that Australia and New Zealand are relatively restricted from other regions, following the pattern that dispersal between southern hemisphere biogeographic regions is a much rarer occurrence when compared to the northern hemisphere.

Notably, armillarioid species are also depauperate in the lower latitudes—an area characterized by largely tropical climate conditions. For example, *A. puiggarii* is the only known mushroom-forming armillarioid species that appears in tropical South America. Similarly, two to three species were broadly found throughout Africa. Perhaps most interesting are the specimens collected in Southeast Asia, which correspond to one species clusters and show no biogeographic connectedness to other regions (Fig. 1), suggesting that further collecting effort here may uncover a novel suite of species that appear to remain unconnected from other biogeographic regions. A temperate bias in collecting efforts undoubtedly plays a role in this finding as this

could certainly be a direct reflection of the disproportionate collecting efforts in temperate regions compared to the tropics. Another explanation could also be a reflection that species in this lineage are better adapted to the cooler and harsher climate in temperate and boreal environments (Koch et al., 2017) or have adapted to temperate woody host species over tropical hosts. Notably, however, it is the tropical rainforests of the Guiana Shield that are the exclusive habitat of the two known sequestrate species of *Guyanagaster* (Henkel et al., 2010; Koch et al., 2017). Given the subhypogeous habit of these sporocarps, it is not implausible that other sequestrate species have yet to be collected and described.

The northern and southern hemispheres have been typically thought to not overlap species distributions (Volk & Burdsall, 1995). While our results are largely consistent with this, there are some exceptions. Only one collection from Argentina is available to support this connection with northern hemisphere species. However, whether this species is a relict from a time when certain *Armillaria* species had a much broader distribution, or a recent migrant to the region on imported plant material, remains unknown. When assessed from a 99% clustering threshold, the specimen from Argentina no longer clusters with the northern hemisphere species, suggesting that it does not represent a recent introduction to the region. We also detected several species dominant throughout the northern hemisphere that were also present in South Africa. Previous phylogenetic analyses showed support at least two introductions of northern hemisphere *Armillaria* species to South Africa which were believed to be introduced centuries ago from potted plants from Europe (Coetzee et al., 2001; Coetzee et al., 2003b; Wingfield et al., 2010). At least one other human-mediated introduction of *Armillaria* from Asia to East Africa has been identified, perhaps introduced to the region on tea plants (Mwenje et al., 2006; Guo et al., 2016).

Of the six and five *Armillaria* species clusters at the ITS1 and ITS2 region, respectively, identified from sequences from African specimens, three of them were also represented in samples from the northern hemisphere. This observation supports previous work that there may be three, or more, independent *Armillaria* introductions to Africa. The other two to three *Armillaria* species clusters from Africa include species only present on that continent. This is consistent with our current understanding of the African diversity of *Armillaria*, and if this remains true and no more putatively native *Armillaria* species are discovered, it represents one of the biogeographic zones with the most species-poor representations of *Armillaria*.

Symbiosis as a driver for distribution and species richness

Another possible contributing factor related to the high species richness in East Asia could be the fact that *Armillaria* species engage in mycorrhizal relationships with mycoheterotrophic plants, notably species of achlorophyllus orchids in the genera *Galeola* and *Gastrodia* (Kikuchi & Yamaji, 2010; Guo et al., 2016). Using the dataset of Guo et al. (2016), four of our *Armillaria* species clusters associate with orchids from an ITS1 perspective, and seven species clusters associate with orchids from an ITS2 perspective (Table 1). None of the species clusters are specific to only orchid associates. However, an interesting hypothesis emerges from this association, as to whether this symbiosis is a driver of species diversification, as some mycorrhizal relationships—particularly ectomycorrhizal—are known to enhance diversification of the symbionts (Koide et al., 2008; Wilson et al., 2017; Sato & Toju, 2019). However, to address this question, we need a more thorough understanding of the geographic extent of this interaction across the global mycoheterotrophic orchid distribution. Such research, particularly in

tropical regions where there are mycoheterotrophic orchids, may also lead to the discovery of new *Armillaria* species.

Another symbiosis that may play a role in shaping the distribution of the *Armillaria* species is that with species of the ascomycete genus *Wynnea*. While it remains unknown what role both of the species in this interaction play, sclerotia produced by *Wynnea* species, regardless of collecting locality, have always been found to contain *Armillaria* hyphae, suggesting a dependency of these two lineages on one another (Xu et al., 2019). Because most of the *Armillaria* sequences from *Wynnea* scloerotia did not pass our filtering thresholds, we independently compared those sequences to our armillarioid species clusters. Armillarioid symbionts of *Wynnea* fell into four species clusters: the clusters 1, 7, 50 and 51 for ITS1 and the clusters 8, 31, 60 and 61 for ITS2. Two of the *Armillaria* symbionts of *Wynnea*, one from West Virginia, USA (species cluster 50 for ITS1 and 60 for ITS2) and the other from Brazil (species cluster 51 for ITS1 and 61 for ITS2), are not conspecific with any of the *Armillaria* species clusters detected in any study and may represent new species. The other two species clusters where *Wynnea* symbionts were detected were those that have large geographic distributions. Even so, for species cluster 1, which had one sporocarp from Central America as a representative of it, there are an additional two *Wynnea* symbionts, from Guatemala and Costa Rica, respectively, that exhibit sequence homology. In species cluster 7, the *Wynnea* symbiont from Brazil extends the geographic distribution of this largely East Asian cluster to tropical South America. One possibility for why some *Wynnea* symbionts were collected in locations that are largely disjunct from where sporocarps were collected may have to do with the specialization of the sclerotium environment in those geographic regions. It is possible that certain armillarioid

species were once more widespread or abundant than they currently are but were outcompeted by better adapted armillarioid species. However, *Wynnea* symbionts may have been buffered from local extinction, particularly in Central and South America, due to the extreme specialization of *Armillaria* species for the *Wynnea* environment. However, more collecting of *Wynnea* sclerotia and armillarioid fruiting bodies, particularly in tropical locales, will be needed to fully test this hypothesis. If this hypothesis is supported, *Wynnea* sclerotia could offer a glimpse into the previous distribution of armillarioid species.

In the *Polyporus umbellatus* interaction, all the *Armillaria* isolates collected by Xing et al. (2017) fell into five and six species clusters when using both ITS1 and ITS2 marker regions, respectively (Table 1). With the ITS1 marker, there is one armillarioid species cluster composed of sequences only known from the association with *P. umbellatus*, suggesting that this may be a source for new *Armillaria* species. This is a similar situation with the putative mycoparasitic interaction between *E. abortivum* and *Armillaria* (Lindner Czederpiltz et al., 2001; Koch & Herr, 2021). In this interaction, the role of symbiosis in driving the distribution and species richness is unclear, often because methods other than the use of ITS region sequencing were used to identify the *Armillaria* associate. Therefore, we do not have a precise understanding of the species that are involved in this interaction. Instead, the distribution of this interaction seems to be driven more so by the distribution of *E. abortivum*, which is restricted to eastern North America and eastern Asia.

Finally, as was previously mentioned, the South American tropics have only a single known mushroom-forming armillarioid species, *A. puiggarii*, but do have two known species of the

sequestrate genus *Guyanagaster*. Phylogenetic and biogeographic analyses show that *Guyanagaster* represents the earliest diverging lineage of the armillarioid clade (Koch et al., 2017). This finding was particularly illuminating because sequestration is a unidirectional process that has never been reversed (Hibbett et al., 1997), suggesting that the ancestor of *Guyanaster* was a gilled mushroom with widespread distribution (Koch et al., 2017). We hypothesized that, because *Guyanagaster* is the only extant lineage, that it is highly specialized for its environment and, as a result, was buffered from extinction unlike its mushroom progenitors (Koch et al., 2017). Further studies suggest its dispersal symbiosis with wood-feeding termites may play a role in its specialization and therefore its persistence (Koch & Aime, 2018; Koch et al., 2021).

CONCLUSION

In the present study, we used publicly available data to address phylogenetic diversity and global species richness of the genus *Armillaria sensu lato*, a group of fungi most notably recognized as mushroom-forming plant pathogens. By mining public data, we estimate that the overall richness of taxa ranges from 50 to 60 species. The highest species richness was observed in Eastern Asia. We also assessed the overlap of species across geographic regions. We proposed numerous hypotheses regarding the drivers of variability in species diversity and richness between different biogeographic regions and along with symbiotic associations with other fungi and host species. Lastly, we show a proof of concept for the use of publicly available sequence data, from individual voucher specimens and from community sequencing efforts, in understanding phylogenetic diversity and species richness in a global context.

Data and code availability.

We used publicly available data as the basis for this study. The public data we used, along with all other associated data that we curated for this study, as well as code for data analysis, can be found at https://github.com/HerrLab/Koch_Global_Armillaria_2021.

Acknowledgements

We want to thank the anonymous reviewers who provided feedback and comments on the initial submission of this manuscript. This work was completed using the Holland Computing Center of the University of Nebraska, which receives support from the Nebraska Research Initiative. This research was directly supported by start-up funding from the University of Nebraska Agricultural Research Division and the University of Nebraska Office of Research and Economic Development. Additionally, JRH acknowledges funding from the US National Air & Space Administration (Grant #80NSSC17K0737), the US National Science Foundation (EPSCoR Grant #1557417), and US National Institute of Justice (Grant #2017-IJ-CX513-0025), all of which indirectly supported this research through the support of research in the laboratory of JRH. All of the funding agencies had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Conflict of Interest Statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Author Contributions

RAK and JRH initiated the work, analyzed the data, designed the figures, and drafted the manuscript.

REFERENCES

Anderson, J. B., and Ullrich, R. C. (1979). Biological species of *Armillaria mellea* in North America. *Mycologia* 71, 402–414.

Banik, M. T., and Burdsall, H. H. (1998). Assessment of compatibility among *Armillaria cepistipes*, *A. sinapina*, and North American biological species X and XI, using culture morphology and molecular biology. *Mycologia* 90, 798–805.

Baumgartner, K., Coetzee, M. P. A., and Hoffmeister, D. (2011). Secrets of the subterranean pathosystem of *Armillaria*. *Mol. Plant Pathol.* 12, 515–534.

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., et al. (2013). Improved software detection and extraction of ITS1 and ITS2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol. Evol.* 4, 914–919.

Bérubé, J. A., and Dessureault, M. (1988). Morphological characterization of *Armillaria ostoyae* and *Armillaria sinapina* sp. nov. *Can. J. Bot.* 66, 2027–2034.

Bérubé, J. A., and Dessureault, M. (1989). Morphological studies of the *Armillaria* mellea complex: two new species, *A. gemina* and *A. calvescens*. *Mycologia* 81, 216–225.

Bruns, T.D., Blackwell, M., Edwards, I., Taylor, A.F., Horton, T., Zhang, N., Kõljalg, U., May, G., Kuyper, T.W., Bever, J.D. and Gilbert, G., 2008. Preserving accuracy in GenBank. *Science*, *319* (5870).

Cha, J. Y., Sung, J. M., and Igarashi, T. (1994). Biological species and morphological characteristics of *Armillaria mellea* complex in Hokkaido: *A. sinapina* and two new species, *A. jezoensis* and *A. singula*. 1994. *Mycoscience* 35, 39–47.

Chillali, M., Wipf, D., Guillaumin, J.-J., Mohammed, C., and Botton, B. (1998). Delineation of the European *Armillaria* species based on the sequences of the internal transcribed spacer (ITS) of ribosomal DNA. *New Phytol.* 138, 553–561.

Coetzee, M. P. A., Wingfield, B. D., Harrington, T. C., Dalevi, D., Coutinho, T. A., and Wingfield, M. J. (2000). Geographical diversity of *Armillaria mellea* s. s. based on phylogenetic analysis. *Mycologia* 92, 105–113.

Coetzee, M. P. A., Wingfield, B. D., Harrington, T. C., Steimel, J., Coutinho, T. A., and Wingfield, M. J. (2001). The root rot fungus *Armillaria mellea* introduced into South Africa by early Dutch settlers. *Mol. Ecol.* 10, 387–396.

Coetzee, M. P. A., Wingfield, B. D., Bloomer, P., Ridley, G. S., and Wingfield, M. J. (2003a). Molecular identification and phylogeny of *Armillaria* isolates from South America and Indo-Malaysia. *Mycologia* 95, 285–293.

Coetzee, M. P. A., Wingfield, B. D., Roux, J., Crous, P. W., Denman, S., and Wingfield, M. J. (2003b). Discovery of two northern hemisphere *Armillaria* species on Proteaceae in South Africa. *Plant Pathol.* 52, 604–612.

Coetzee, M. P. A., Bloomer, P., Wingfield, M. J., and Wingfield, B. D. (2011). Paleogene radiation of a plant pathogenic mushroom. *PLoS One* 6, e28545.

Coetzee, M. P. A., Wingfield, B. D., Zhao, J., van Coller, S. J., and Wingfield, M. J. (2015). Phylogenetic relationships among biological species of *Armillaria* from China. *Mycoscience* 56, 530–541.

Coetzee, M. P. A., Wingfield, B. D., and Wingfield, M. J. (2018). *Armillaria* Root-Rot pathogens: species boundaries and global distribution. *Pathogens* 7, 83.

Darmono, T. W., Burdsall, H. H., and Volk, T. J. (1992). Interfertility among isolates of *Armillaria tabescens* in North America. *Sydowia* 42, 105–116.

Dunne, C. P., Glen, M., Tommerup, I. C., Shearer, B. L., and Hardy, G. E. S. J. (2002). Sequence variation in the rDNA ITS of Australian *Armillaria* species and intra-specific variation in *A. luteobubalina*. *Australas. Plant Pathol.* 31, 241–251.

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461.

Federhen, S. (2012). The NCBI Taxonomy database. *Nucleic Acids Res.* 40, D136–D143.

Garraway, M. O., Hütterman, A., and Wargo, P. M. (1991). “Ontogeny and physiology,” in Armillaria Root Disease Agriculture Handbook 691, ed. C. G. Shaw and G. A. Kile (Washington DC: USDA Forest Service) 21–47.

Guillaumin, J.-J., Mohammed, C., Anselmi, N., Courtecuisse, R., Gregory, S. C., Holdenrieder, O., et al. (1993). Geographical distribution and ecology of the *Armillaria* species in western Europe. *Eur. J. Plant Pathol.* 23, 321–341.

Guo, S., and Xu, J. (1992). Nutrient source of sclerotia of *Grifola umbellata* and its relationship to *Armillaria mellea*. *Acta Bot. Sin.* 34, 576–580.

Guo, T., Wang, H. C., Xue, W. C., Zhao, J., and Yang, Z. L. (2016). Phylogenetic analyses of *Armillaria* reveal at least 15 phylogenetic lineages in China, seven of which are associated with cultivated *Gastrodia elata*. *PLoS One* 11, e0154794.

Haelewaters, D., Dirks, A. C., Kappler, L. A., Mitchell, J. K., Quijada, L., Vandegrift, R., et al. (2018). A preliminary checklist of fungi at Boston Harbor Islands. *Northeast. Nat.* 25, 45–76.

Hawksworth, D. L., and Lücking, R. (2017). Fungal diversity revisited: 2.2 to 3.8 million species. *Microbiol. Spectr.* 5, FUNK-0052-2016.

Henkel, T. W., Smith, M. E., and Aime, M. C. (2010). *Guyanagaster*, a new wood-decaying sequestrate fungal genus related to *Armillaria* (Physalacriaceae, Agaricales, Basidiomycota). *Am. J. Bot.* 97, 1474–1484.

Herr, J.R., Opik, M. and Hibbett, D.S., 2015. Towards the Unification of Sequence-Based Classification and Sequence-Based Identification of Host-Associated Microorganisms. *New Phytologist*, *205*(1), pp. 27-31.

Hibbett, D. S., Pine, E. M., Langer, E., Langer, G., and Donoghue, M. J. (1997). Evolution of gilled mushrooms and puffballs inferred from ribosomal DNA sequences. *Proc. Natl. Acad. Sci. U.S.A.* 94, 12002–12006.

Hibbett, D., Abarenkov, K., Kõljalg, U., Öpik, M., Chai, B., Cole, J., Wang, Q., Crous, P., Robert, V., Helgason, T. et al.., 2016. Sequence-based classification and identification of Fungi. *Mycologia*, *108*(6), pp.1049-1068.

Hood, I. A., and Ramsfield, T. D. (2016). *Armillaria aotearoa* species nova. *N. Z. J. For. Sci.* 46, 2.

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780.

Kikuchi, G., and Yamaji, H. (2010). Identification of *Armillaria* species associated with *Polyporus umbellatus* using ITS sequences of nuclear ribosomal DNA. *Mycoscience* 51, 366–372.

Kile, G. A., McDonald, G. I., and Byler, J. W. (1991). Ecology and disease in natural forests," in Armillaria Root Disease Agriculture Handbook 691, ed. C. G. Shaw and G. A. Kile (Washington DC: USDA Forest Service) 102–121.

Klopfenstein, N. B., Stewart, J. E., Ota, Y., Hanna, J. W., Richardson, B. A., Ross-Davis, A. L., et al. (2017). Insights into the phylogeny of Northern Hemisphere Armillaria: Neighbor-network and Bayesian analyses of translation elongation factor 1-α gene sequences. *Mycologia* 109, 75–91.

Koch, R. A., and Aime, M. C. (2018). Population structure of *Guyanagaster necrorhizus* supports termite dispersal for this enigmatic fungus. *Mol. Ecol.* 27, 2667–2679.

Koch, R. A., and Herr, J. R. In review. Transcriptomics reveals the mycoparasitic strategy of the mushroom *Entoloma abortivum* on species of the mushroom *Armillaria*. Submitted to mSystems. On bioRxiv as https://doi.org/10.1101/2021.04.30.442184

Koch, R. A., Wilson, A. W., Séné, O., Henkel, T. W., and Aime, M. C. (2017). Resolved phylogeny and biogeography of the root pathogen *Armillaria* and its gasteroid relative, *Guyanagaster*. *BMC Evol. Biol.* 17, 33.

Koch, R. A., Yoon, G. M., Arayal, U. K., Lail, K., Amirebrahimi, M., LaButti, K., et al. (2021). Symbiotic nitrogen fixation in the reproductive structures of a basidiomycete fungus. Current Biology, accepted.

Koide, R.T., Sharda, J.N., Herr, J.R. and Malcolm, G.M., 2008. Ectomycorrhizal fungi and the biotrophy–saprotrophy continuum. *New Phytologist*, *178*(2), pp.230-233.

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., and Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* 35, 4453–4455.

Kuske, C. R., Hesse, C. N., Challacombe, J. F., Cullen, D., Herr, J. R., Mueller, R. C., et al. (2015). Prospects and challenges for fungal metatranscriptomics of complex communities. *Fungal Ecol.* 14, 133–137.

Leinonen, R., Sugawara, H., Shumway, M., and International Nucleotide Sequence Database Collaboration. (2010). The sequence read archive. *Nucleic Acids Res.* 39, D19–D21.

Liimatainen, K., Niskanen, T., Dima, B., Kytövuori, I., Ammirati, J. F., and Frøslev, T. G. (2014). The largest type study of Agaricales species to date: bringing identification and nomenclature of Phlegmacium (Cortinarius) into the DNA era. *Persoonia* 33, 98–140.

Lindner Czederpiltz, D. L., Volk, T. J., and Burdsall, H. H., Jr. (2001). Field observations and inoculation experiments to determine the nature of the carpophoroids associated with *Entoloma abortivum* and *Armillaria*. *Mycologia* 93, 841–851.

Maphosa, L., Wingfield, B. D., Coetzee, M. P. A., Mwenje, E., and Wingfield, M. J. (2006). Phylogenetic relationships among *Armillaria* species inferred from partial elongation factor 1-alpha DNA sequence data. *Australas. Plant Pathol.* 35, 513–520.

Mwenje, E., Wingfield, B. D., Coetzee, M. P. A., Nemato, H., and Wingfield, M. J. (2006). *Armillaria* species on tea in Kenya identified using isozyme and DNA sequence comparisons. *Plant Pathol.* 55, 343–350.

Nilsson, R. H., Larsson, K.-H., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., et al. (2018). The UNTE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* 47, D259–D264.

Oliveira, J. J. S., Moncalvo, J.-M., Margaritescu, S., and Capelari, M. (2020). A morphological and phylogenetic evaluation of *Marasmius* sect. *Globulares* (Globulares-Sicci complex) with nine new taxa from the Neotropical Atlantic Forest. *Persoonia* 44, 240–277.

Ota, Y., Matsushita, N., Nagasawa, E., Terashita, T., Fukuda, K., and Suzuki, K. (1998). Biological species of *Armillaria* in Japan. *Plant Dis.* 82, 537–543.

Pildain, M. B., Coetzee, M. P. A., Rajchenberg, M., Petersen, R. H., Wingfield, M. J., and Wingfield, B. D. (2009). Molecular phylogeny of *Armillaria* from the Patagonian Andes. *Mycol. Prog.* 8, 181–194.

Qin, G. F., Zhao, J., and Korhonen, K. (2007). A study on intersterility groups of *Armillaria* in China. *Mycologia* 99, 430–441.

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rognes, T., Flouri, T., Nichols, B., Quince, C., and Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584.

Sato, H., and Toju, H. (2019). Timing of evolutionary innovation: scenarios of evolutionary diversification in a species-rich fungal clade, Boletales. *New Phytol.* 222, 1924–1935.

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., Chen, W., et al. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for *Fungi. Proc. Natl. Acad. Sci. U.S.A.* 109, 6241–6246.

Sipos, G., Prasanna, A.N., Walter, M.C., O'Connor, E., Bálint, B., Krizsán, K., Kiss, B., Hess, J., Varga, T., Slot, J. and Riley, R., 2017. Genome expansion and lineage-specific genetic innovations in the forest pathogenic fungi *Armillaria*. *Nature ecology & evolution*, *1*(12), pp.1931-1941.

Smith, M. L., Bruhn, J. N., and Anderson, J. B. (1992). The fungus *Armillaria bulbosa* is among the largest and oldest living organisms. *Nature* 356, 428–431.

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., et al. (2014). Global diversity and geography of soil fungi. *Science* 346, 1256688.

Truong, C., Mujic, A. B., Healy, R., Kuhar, F., Furci, G., Torres, D., et al. (2017). How to know the fungi: combining field inventories and DNA-barcoding to document fungal diversity. *New Phytol.* 214, 913–919.

Tsykun, T., Rigling, D., and Prospero, S. (2013). A new multilocus approach for a reliable DNA-based identification of *Armillaria* species. *Mycologia* 105, 1059–1076.

Volk, T. J., and Burdsall, H. H., Jr. (1995). A Nomenclatural Study of *Armillaria* and *Armillariella* Species. Oslo: Synopsis Fungorum.

Vu, D., Groenewald, M., de Vries, M., Gehrmann, T., Stielow, B., Eberhardt, U., et al. (2019). Large-scale generation and analysis of filamentous fungal DNA barcodes boosts coverage for kingdom fungi and reveals thresholds for fungal species and higher taxon delimation. *Stud. Mycol.* 92, 135–154.

Watling, R., Kile, G. A., and Burdsall, H. H., Jr. (1991). “Nomenclature, taxonomy, and identification,” in Armillaria Root Disease Agriculture Handbook 691, ed. C. G. Shaw and G. A. Kile (Washington DC: USDA Forest Service) 1–9.

Wilson, A. W., Hosaka, K., and Mueller, G. M. (2017). Evolution of ectomycorrhizas as a driver of diversification and biogeographic patterns in the model mycorrhizal mushroom genus *Laccaria*. *New Phytol.* 213, 1862–1873.

Wingfield, M. J., Coetzee, M. P. A., Crous, P. W., Six, D., and Wingfield, B. D. (2010). Fungal phoenix rising from the ashes? *IMA Fungus* 1, 149–153.

Wright, E. S. (2016). Using DECIPHER v2.0 to analyze big biological sequence data in R. *R J.* 8, 352–359.

Xing, X., Men, J., and Guo, S. (2017). Phylogenetic constrains on *Polyporus umbellatus*-*Armillaria* associations. *Sci. Rep.* 7, 4226.

Xu, F., LoBuglio, K. F., and Pfister, D. H. (2019). On the co-occurrence of species of *Wynnea* (*Ascomycota, Pezizales, Sarcoscyphaceae*) and *Armillaria* (*Basidiomycota, Agaricales, Physalacriaceae*). *Fungal Syst. Evol.* 4, 1–12.

Yu, G., Tsan-Yuk Lam, T., Zhu, H., and Guan, Y. 2018. Two methods for mapping and visualizing associated data on phylogeny using Ggtree. *Mol. Biol. Evol.* 35, 3041–3043.

Zhao, J., Dai, Y. C., Qin, G. F., Wei, Y. L., Wang, H. C., and Yuan, H. S. (2008). New biological species of *Armillaria* from China. *Mycosystema* 27, 156–170.

Zolciak, A., Bouteville, R.-J., Tourvieille, J., Roeckel-Drevet, P., Nicolas, P., and Guillaumin, J.-J. (1997). Occurrence of *Armillaria ectypa* (Fr.) Lamoure in peat bogs of the Auvergne—the reproduction system of the species. *Cryptogam. Mycol.* 18, 299–313.

FIGURE LEGENDS

Fig. 1. Connectedness of biogeographic regions by shared armillarioid species clusters. The width of the lines and diameter of circles are proportional to the number of connections and number of species, respectively. The blue line corresponds to ITS1 and the magenta line to ITS2. The numbers within the circles indicate the number of armillarioid species clusters detected in each region, with the top number referring to the number of species clusters at ITS1 and the bottom number at ITS2.

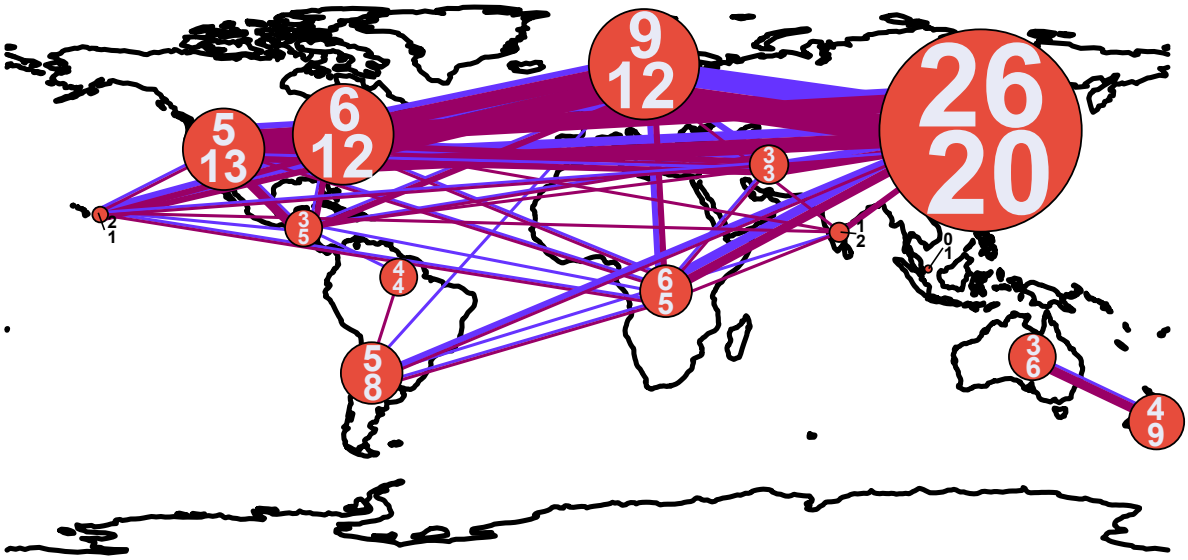
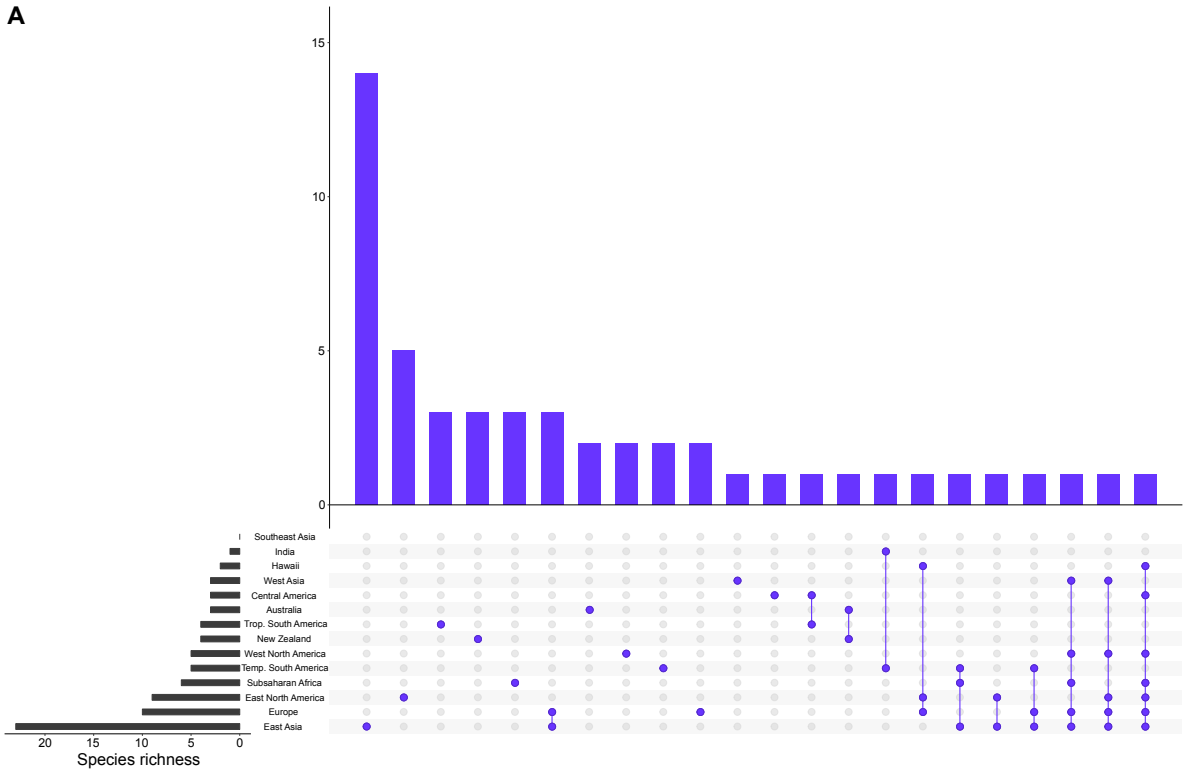

9
12
26
20
6
12
5
13
3
3
2
1
3
5
1
2
0
1
4
4
6
5
5
8
3
6
4
9

Fig. 2. Species richness in each of the biogeographic regions. Bar graphs showing the total richness of our species clusters in each biogeographic region and the combination of biogeographic regions. (A) ITS1 and (B) ITS2. The species richness for each vertical bar corresponds to the number of species that are solely found in the biogeographic regions corresponding to the colored circle(s) below. The species richness for the horizontal insets correspond to the total species richness within each biogeographic region.


A
15
10
5
0
Southeast Asia
India
Hawaii
West Asia
Central America
Australia
Trop. South America
New Zealand
West North America
Temp. South America
Subsaharan Africa
East North America
Europe
East Asia
20
15
10
5
0
Species richness

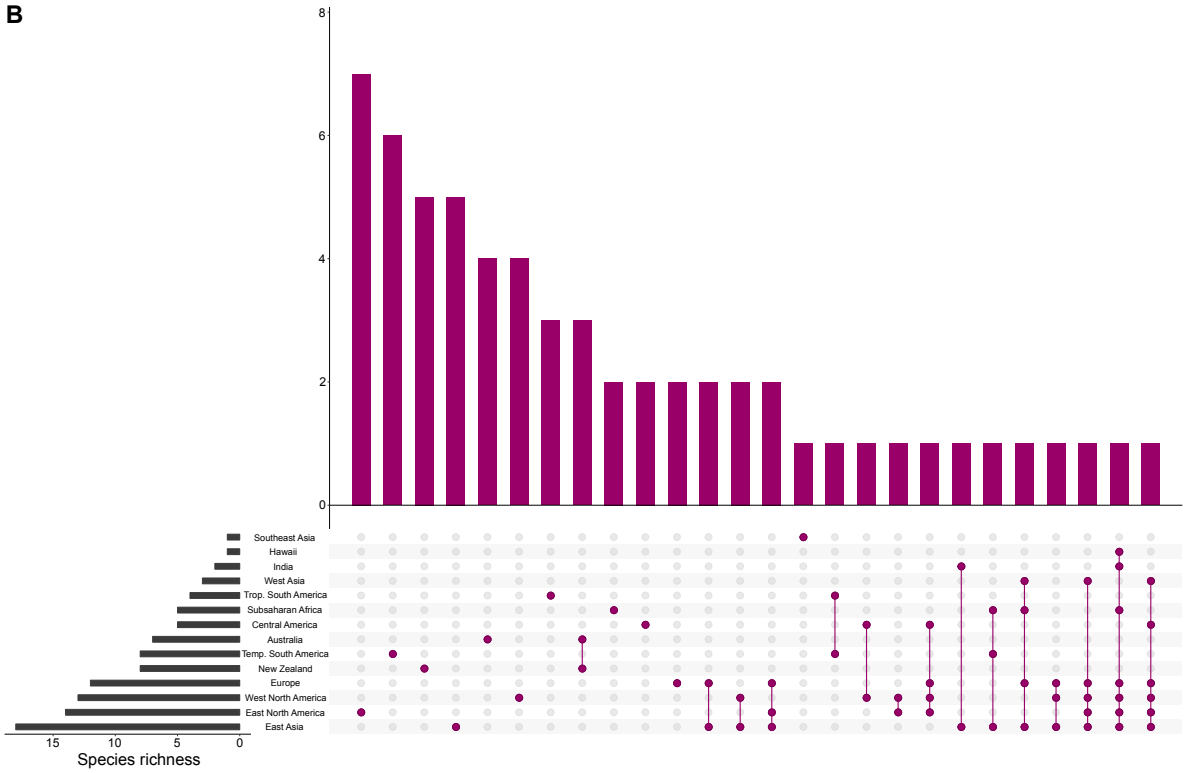

B
8
6
4
2
0
Southeast Asia
Hawaii
India
West Asia
Trop. South America
Subsaharan Africa
Central America
Australia
Temp. South America
New Zealand
Europe
West North America
East North America
East Asia
15
10
5
0
Species richness

Fig. 3. Phylogenetic trees illustrating the number of sequences that correspond to each species cluster. Maximum likelihood phylogeny constructed from the full length ITS region using one exemplar sequence per species cluster. Two species of *Guyanagaster*, sister to all other armillarioid species, were used to root this phylogeny. Well-supported branches (greater than 70% bootstrap support) are thickened. The color of the corresponding circles corresponds to the biogeographic region from which the voucher material producing the sequences was collected. The size of the circles is scaled by the number of sequences that were grouped into a particular species cluster from a particular region. (A) Phylogeny of species clusters as determined from analysis of ITS1 and (B) phylogeny of species clusters as determined from analysis of ITS2.

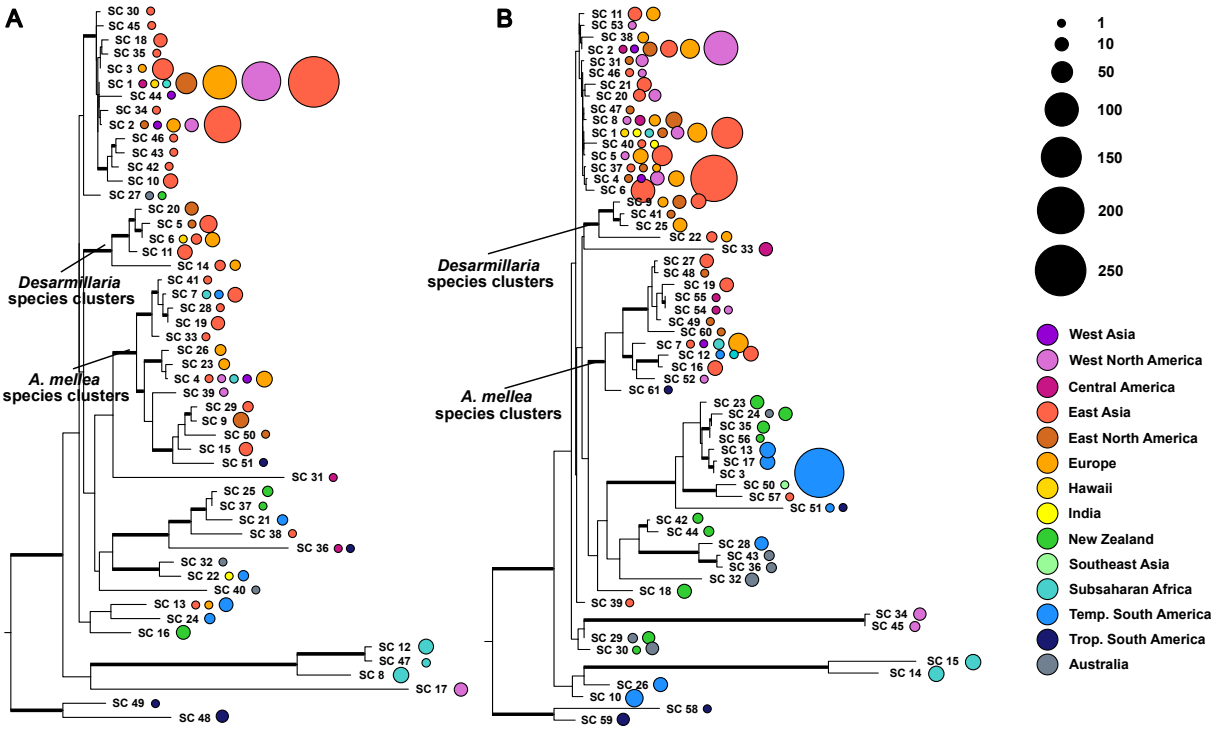

A
B
Desarmillaria species clusters
A. mellea species clusters
1
10
50
100
150
200
250
West Asia
West North America
Central America
East Asia
East North America
Europe
Hawaii
India
New Zealand
Southeast Asia
Subsaharan Africa
Temp. South America
Trop. South America
Australia

Table 1. Armillarioid sequences (and corresponding GenBank accessions and location metadata) from tissue engaging in symbiotic associations.

	Species cluster	
***Wynnea* symbionts**	**ITS1**	**ITS2**
MK271357: Pennsylvania, USA	1	8
MK281621: China	1	8
MK281620: Guatemala	1	8
MK271358: Costa Rica	1	8
MK271360: New York, USA	1	8
MK271359: New York, USA	1	8
MK271361: West Virginia, USA	1	8
MK599142: China	1	8
MK271362: Brazil	7	12
MK281622: Japan	7	12
MK281623: West Virginia, USA	50	60
MK281619: Brazil	51	61
***Polyporus* symbionts**	**ITS1**	**ITS2**
KY389148: China	1	1
KY389149: China	1	1
KY389150: China	1	1
KY389170: China	1	1
KY389172: China	1	1
KY389174: China	1	1
KY389176: China	1	1
KY389189: China	1	1
KY389154: China	1	4
KY389167: China	1	4
KY389168: China	1	4
KY389169: China	1	4
KY389171: China	1	4
KY389184: China	1	4
KY389185: China	1	4
KY389191: China	1	4
KY389192: China	1	4
KY389160: China	1	5
KY389161: China	1	5
KY389162: China	1	5
KY389163: China	1	5
KY389164: China	1	5
KY389165: China	1	5
KY389166: China	1	5
KY389159: China	1	6
KY389178: China	1	6
KY389182: China	1	6
KY389183: China	1	6
KY389155: China	2	2

KY389147: China	2	4
KY389153: China	2	4
KY389177: China	2	4
KY389180: China	2	4
KY389181: China	2	4
KY389186: China	2	4
KY389187: China	2	4
KY389179: China	2	6
KY389188: China	2	6
KY389193: China	2	6
KY389190: China	3	2
KY389152: China	3	4
KY389157: China	3	4
KY389158: China	3	4
KY389151: China	3	6
KY389173: China	3	6
KY389156: China	35	3
KY389175: China	47	1

Orchid symbionts	**ITS1**	**ITS2**
KT822259: China	1	1
KT822279: China	1	1
KT822297: China	1	1
KT822314: China	1	1
KT822334: China	1	1
KT822336: China	1	1
KT822286: China	1	4
KT822303: China	1	4
KT822329: China	1	4
KT822260: China	1	5
KT822301: China	1	6
KT822270: China	1	20
KT822322: China	2	46
KT822296: China	3	1
KT822264: China	3	6
KT822268: China	3	6
KT822273: China	3	6
KT822284: China	3	6
KT822285: China	3	6
KT822289: China	3	6
KT822275: China	3	21
KT822276: China	3	21
KT822302: China	30	1