Title

Mycoparasites, Gut Dwellers, and Saprotrophs: Phylogenomic Reconstructions and Comparative Analyses of Kickxellomycotina Fungi

Journal

Genome Biology and Evolution, 15(1)

ISSN

1759-6653

Authors

Reynolds, Nicole K
Stajich, Jason E
Benny, Gerald L
et al.

Publication Date

2023-01-04

DOI

10.1093/gbe/evac185

Research Article

Title: Mycoparasites, gut dwellers, and saprotrophs: Phylogenomic reconstructions and comparative analyses of Kickxellomycotina fungi

Nicole K. Reynolds[1*], Jason E. Stajich[2], Gerald L. Benny[1], Kerrie Barry[3], Stephen Mondo[3], Kurt LaButti[3], Anna Lipzen[3], Chris Daum[3], Igor V. Grigoriev[3,4], Hsiao-Man Ho[5], Pedro W. Crous[6], Joseph W. Spatafora[7], Matthew E. Smith[1]

[1] University of Florida, Department of Plant Pathology, Gainesville, Florida USA 32611

[2] *Department of Microbiology & Plant Pathology and Institute for Integrative Genome Biology, University of California–Riverside, Riverside, California USA 92521*

[3] *U.S. Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, California USA 94720*

[4] *Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, California USA 94720*

[5] *Department of Science Education, National Taipei University of Education, 134, Section 2, Heping E. Road, Taipei 106, Taiwan*

[6] *Westerdijk Fungal Biodiversity Institute, Uppsalalaan 8, 3584 CT, Utrecht, The Netherlands*

[7] *Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon USA 97331*

**Corresponding author: nr299@cornell.edu*

Abstract

Improved sequencing technologies have profoundly altered global views of fungal diversity and evolution. High throughput sequencing methods are critical for studying fungi due to the cryptic, symbiotic nature of many species, particularly those that are difficult to culture. However, the low coverage genome sequencing (LCGS) approach to phylogenomic inference has not been widely applied to fungi. Here we analyzed 171 Kickxellomycotina fungi using LCGS methods to obtain hundreds of marker genes for robust phylogenomic reconstruction. Additionally, we mined our LCGS data for a set of nine rDNA and protein coding genes to enable analyses across species for which no LCGS data were obtained. The main goals of this study were to: 1) evaluate the quality and utility of LCGS data for both phylogenetic reconstruction and functional annotation, 2) test relationships among clades of Kickxellomycotina, and 3) perform comparative functional analyses between clades to gain insight into putative trophic modes. In opposition to previous studies, our nine-gene analyses support two clades of arthropod gut dwelling species and suggest a possible single evolutionary event leading to this symbiotic lifestyle. Furthermore, we resolve the mycoparasitic Dimargaritales as the earliest diverging clade in the subphylum and find four major clades of *Coemansia* species. Finally, functional analyses illustrate clear variation in predicted carbohydrate active enzymes and secondary metabolites (SM) based on ecology, i.e., biotroph vs. saprotroph. Saprotrophic Kickxellales broadly lack many known pectinase families compared to saprotrophic Mucoromycota and are depauperate for SM but are enriched for chitinases compared to biotrophic taxa in Zoopagomycota.

Significance statement

Environmental sequencing efforts indicate that much undiscovered fungal diversity lies within early diverging clades, but these species are often cryptic and challenging to collect. Molecular tools are important for understanding the biology of these fungi, but due to difficulties associated with culturing, obtaining adequate DNA can be a limiting factor. Our analyses of Kickxellomycotina fungal genomes show that useful genetic data can be obtained from samples with as little as 600ng of DNA. These data can be used as input for phylogenomic inference as well as preliminary, clade-based comparisons of functional annotations. Our results illustrate the utility of low coverage genomic data for uncovering cryptic species diversity, generating new evolutionary hypotheses, and inferring potential ecological roles.

Keywords: Automatic Assembly for the Fungi, Funannotate, low coverage phylogenomics, PHYling, trichomycetes

INTRODUCTION

Understanding fungal evolution is challenging due to the paucity of fungal fossils, the cryptic and microscopic nature of many fungi, and the large number of estimated fungal taxa (2.2 to 3.8 million species – Hawksworth & Lücking 2017). Even the most obvious macroscopic form, the fungal fruiting body, is a visible yet ephemeral portion of a lifecycle that mainly takes place below-ground or in other hidden habitats (e.g., inside another organism as a symbiont). DNA sequencing technology has transformed fungal systematics and simultaneously increased our estimates of fungal diversity. Molecular

phylogenetic studies have also revealed prolific convergent evolution in many ecological niches (e.g., mycoparasitism, lichenization, ectomycorrhizal associations) (Ahrendt et al. 2018; Chang et al. 2018; Grube & Wedin 2016) and physical structures (e.g., hypogeous truffle-like forms, losses of flagella) (Bonito et al. 2013; Galindo et al. 2021). Fungal genome sequencing efforts like the 1,000 Fungal Genomes Project (1000.fungalgenomes.org) and the Zygomycetes Genealogy of Life Project (zygolife.org) have made great strides towards an inclusive fungal phylogeny, particularly by increasing taxon sampling among the early diverging fungal lineages. Species of zygomycetes are the earliest diverging fungi to evolve a hyphal growth form and were previously classified within a single phylum (Zygomycota), but phylogenomic analyses demonstrated that they constitute two monophyletic phyla: Mucoromycota and Zoopagomycota (Spatafora et al. 2016), although other phylum-level relationships have been proposed (e.g., Li et al. 2021; Strassert & Monaghan 2022). Mucoromycota comprise primarily plant-associated and saprotrophic species whereas Zoopagomycota includes mostly parasites of fungi and animals. Zoopagomycota species are less frequently collected, and many are difficult to culture and maintain in the lab, so members of this group have been understudied and less frequently included in phylogenetic analyses (Benny et al. 2016). However, technological advances such as single-cell genome sequencing have generated additional data for challenging parasitic taxa (Ahrendt et al. 2018; Davis et al. 2019). Genome-scale data have also facilitated new insights into the biology of Zoopagomycota fungi by elucidating metabolic pathways and secondary metabolite production (Ahrendt et al. 2018; Tabima et al. 2020), the evolution of plant cell wall degrading enzymes (Chang et al. 2015), and horizontal gene

transfer events (Tabima et al. 2020; Wang et al. 2016). Zoopagomycota are further divided into 3 subphyla: Entomophthoromycotina (insect parasites), Zoopagomycotina (parasites of fungi and small animals such as nematodes) and Kickxellomycotina (several trophic strategies, see below) (Hibbett et al. 2007; Spatafora et al. 2016). Phylogenomic analyses incorporating these taxa found that Kickxellomycotina and Zoopagomycotina are both monophyletic lineages whereas Entomophthoromycotina is polyphyletic (Davis et al. 2019; Li et al. 2021). However, most studies have included only a few species from each lineage.

Kickxellomycotina are an interesting target for evolutionary studies because the group includes a monophyletic lineage of mycoparasites (Dimargaritales) and a mixture of non-monophyletic saprotrophs (Kickxellales, Ramicandelaberales, Spiromycetales) and arthropod gut symbiont lineages ("trichomycetes") (Asellariales, Harpellales, Orphellales, Barbatosporales) (Doweld 2014; Tretter et al. 2014; White et al. 2018). Most Kickxellales species are apparently rare in the environment and some taxa have been reported only once or a few times (e.g., *Dipsacomyces acuminosporus, Mycoëmilia yatsukahoi*, *Spirodactylon aureum*) (Benjamin 1959, 1961; Kurihara et al. 2004). The exception is the genus *Coemansia*, which is the most commonly encountered genus of Kickxellales from soil and dung samples (Benny et al. 2016). Species of Dimargaritales parasitize fungal hosts in the Mucoromycota (e.g., *Cokeromyces*) and Ascomycota (e.g., *Chaetomium*) via a specialized apparatus that includes a penetration cell (appressorium) and a biotrophic absorption cell (haustorium) (Benjamin 1959, 1961). Dimargaritales species are almost exclusively found on dung rather than soil and are less common than Kickxellales species. Species of Harpellales

live as commensals in the digestive tracts of immature aquatic insects where they attach to the host via a specialized holdfast cell or glue-like material (Lichtwardt et al. 2001). Asellariales, Orphellales, and Barbatosporales are generally ecologically and morphologically similar to Harpellales. However, these fungi inhabit different hosts, have some unique morphological characters, and all four arthropod-associated groups are hypothesized to represent three independent evolutionary events leading to arthropod symbiosis (Tretter et al. 2014; Valle & Cafaro 2008; White et al. 2006; White et al. 2018).

Phylogenetic studies based on rDNA or multigene datasets support the sister relationship between Kickxellales and Orphellales and also between Harpellales and Asellariales, whereas Dimargaritales and Ramicandelaberales have been resolved in the earliest-diverging clade (Tretter et al. 2014; White et al. 2006, 2018). However, phylogenetic placement of Dimargaritales has been hampered by limited taxon sampling in genomic studies (Li et al. 2021; Ahrendt et al. 2018; Davis et al. 2019) and topology conflicts in multigene studies (Tretter et al. 2014). Phylogenies of Kickxellales based on rDNA data have also found polyphyly among genera, with the monotypic genera *Kickxella* (the type genus of the order) and *Spirodactylon* nested within the speciose genus *Coemansia* (Chuang et al. 2018; White et al. 2018). Additionally, rDNA data were not able to fully resolve relationships among some *Coemansia* taxa (Chuang et al. 2018).

Recent comparative genomic analyses indicate that species within different lineages of Kickxellomycotina have vastly different genome sizes and predicted secondary metabolites (Ahrendt et al. 2018; Tabima et al. 2020). Kickxellales and the

mycoparasites in Dimargaritales have notably smaller estimated genome sizes (~20-30 Mb) than the insect-symbiotic Harpellales (~70-100 Mb). *Dimargaris cristalligena* (Dimargaritales), *Linderina pennispora* (Kickxellales), and *Martensiomyces pterosporus* (Kickxellaes) had more predicted secondary metabolites than all other analyzed Zoopagomycota except for species of *Basidiobolus* (Entomophthoromycotina). In contrast, *Coemanisa mojavensis*, *C. reversa*, and *C. spiralis* (Kickxellales) had among the fewest secondary metabolites (Tabima et al. 2020). The predicted secondary metabolite classes were also variable among the different lineages, with *D. cristalligena* having more non-ribosomal peptide synthetases and Kickxellales having more polyketide synthases (Tabima et al. 2020). Ahrendt et al. (2018) analyzed genomes from various mycoparasites and found that some subtilase genes of *D. cristalligena* formed a distinct group, while others overlapped with other Zoopagomycota sequences. Similar to other mycoparasites, *D. cristalligena* lacked several enzymes usually present in the metabolic pathways of related saprobic fungi (e.g., enzymes in the thiamine synthesis and sulfate reduction pathways). Finally, Zoopagomycota fungi are generally assumed to be haploid, but single nucleotide polymorphism comparisons between single-cell genomes versus genomes generated from multiple cells indicated that *D. cristalligena* is likely not haploid (Ahrendt et al. 2018). These findings point to the uniqueness of *D. cristalligena* relative to other Kickxellomycotina and illustrate various genomic features among the different lineages which might be related to their diverse ecological niches.

While full coverage reference sequences paired with transcriptomic data are the ultimate goal for genome sequencing, this requires large amounts of pure, high-quality

DNA and RNA. Single cell genomics methods are a useful alternative for unculturable or fastidious species, but this approach is relatively expensive, sensitive to contamination, and prone to various sequencing artifacts (Gawad et al. 2016; Pinard et al. 2006; Sabina and Leamon 2016). An alternative approach is low coverage genome sequencing (LCGS) using a high throughput, short read technology, such as Illumina (e.g., Zhang et al. 2019; Liimatainen et al. 2022). Here we define LCGS as sequencing based on short read technology without a corresponding transcriptome. We do not include the number of reads per base in our definition due to the small expected genome size of our species; therefore we anticipate high reads per base coverage that may not correspond to a high quality assembly (e.g., low BUSCO complete scores, large number of contigs). The quality of assembly can also depend on the repetitive sequence content, which is quite low for the Kickellomycotina (2-6%) (Ahrendt et al. 2018; Amses et al. 2022; Chang et al. 2015; Chang et al. 2022; Mondo et al. 2017; Wang et al. 2016). There are several conditions under which LGCS would be a more appropriate choice than other genome sequencing methods, including: 1) phylogenetic reconstruction is the main goal of data analysis, 2) the genome sizes of the target species are small and high coverage can be achieved with short read sequencing, and 3) many isolates will be sequenced (and therefore cost is an important consideration). The LGCS method has proven useful in phylogenomic studies of another zygomycete lineage, Mortierellaceae (Mucoromycota) (Vandepol et al. 2020). Most Mortierellaceae species have a ~40 Mb genome and conflicts between morphology and molecular phylogenies had caused taxonomic confusion within the group (Petkovits et al. 2011). A combination of LGCS and multigene data from 318 Mortierellaceae isolates was used to

resolve the phylogeny and disentangle several longstanding taxonomic problems (Vandepol et al. 2020).

Kickxellomycotina are an ideal candidate for phylogenomic analyses using LGCS. Most saprotrophic Kickxellomycotina species have genomes of ~15-25 Mb, many species can be grown in pure or dual cultures, and relationships within and between several major lineages remain unresolved. We used Illumina sequencing to generate LGCS data for 171 Kickxellomycotina isolates (including 158 *Coemansia* species) and we also used PacBio to generate a full coverage, reference genome (including a transcriptome) for *Pinnaticoemansia coronantispora* CBS 131509 (Kickxellales). We had several hypotheses regarding methodology, phylogenetic topology, and functional annotation. As far as methodology, we hypothesized: 1) we would achieve high coverage due to small predicted genomes thus allowing us to retrieve hundreds of marker genes for phylogenetic reconstruction and preliminary functional analyses, and 2) reconstruction using different marker sets (i.e., BUSCO vs. Orthofinder) and analysis methods (i.e., maximum likelihood vs. coalescent-based) would yield the same topology due to the high number of orthologs recovered. Regarding phylogenetic reconstruction, we hypothesized: 1) *Coemansia* would be polyphyletic (with *Kickxella* and *Spirodactylon* nested within), 2) species complexes within *Coemansia* and the placement of Dimargaritales and Ramicandelaberales would be resolved, and 3) arthropod associated taxa would remain as separate monophyletic clades (as found previously). For functional annotation we hypothesized: 1) all species of Dimargaritales would have greater numbers of predicted secondary metabolites and proteases than other Kickxellomycotina, 2) *Coemansia* species would be depauperate

in these enzyme predictions, and 3) the enzymatic profiles of *Coemansia* species would be consistent with a saprotrophic lifestyle (i.e., higher proportion of carbohydrate active enzymes than proteases).

RESULTS

Nine-gene phylogeny—To create a dataset that could incorporate sequences from species lacking genome data, we mined our LGCS data (SUPP TABLE 1) for a set of eight marker genes (18S, 28S, β tubulin, EF1α, MCM7, RPB1, RPB2, and TSR1) used by Tretter et al. (2014) (plus one additional gene, actin) using the aTRAM (Allen et al. 2015) and PHYling pipelines (Stajich, http://github.com/stajichlab/PHYling _unified). The PHYling pipeline recovered most of the nine target genes from the majority of isolates in the LGCS dataset. Of the 171 analyzed genomes, PHYling identified RPB1 from the most isolates (n=171) and EF1α from the fewest (n=147). Using the aTRAM pipeline, 18S and 28S rDNA data were recovered from all *Coemansia* and *Kickxella* isolates, but not from the Dimargaritales. Only host genes were recovered from Dimargaritales samples, so no rDNA data were included for those isolates. After adding sequences downloaded from GenBank, the final alignment contained 227 representatives of Kickxellomycotina. Once unaligned regions were excluded, the final concatenated alignment contained 5,571 characters and 24.1% missing data. A few taxa (e.g., *Myconymphaea yatsukahoi*) only had rDNA sequences available. The best Maximum Likelihood (ML) tree is shown in FIG 1. Four main clades of *Coemansia* were recovered and we refer to them as *Coemansia*, Kickxellales1, Kickxellales2, and *Kickxella*. The subtending nodes of the *Coemansia*, Kickxellales1, and Kickxellales2 clades were unsupported. *Spirodactylon aureum* NRRL 2810 was placed within the *Coemansia*

clade (sister to *Coemansia* sp. RSA 552), recapitulating the topology of the rDNA tree from Chuang et al. (2018). The isopod-associated *Asellaria ligiae* (Asellariales) and blackfly-associated *Barbatospora ambicaudata* (Barbatosporales) were placed as early diverging branches within the insect-associated Harpellales clade and the Harpellales + Asellariales + Barbatosporales were recovered as sister to the coprophilous Spiromycetales. The stonefly-associated Orphellales were sister to the Harpellales + Spiromycetales. The mycoparasitic Dimargaritales were the earliest diverging lineage in the Kickxellomycotina, followed by the putatively saprotrophic Ramicandelaberales as sister to all other orders. Although the Dimargaritales were recovered as a monophyletic group, *Dimargaris* and *Dispira* were polyphyletic. Taxonomic emendations are in the supplemental files.

Genome assembly—Low coverage genome data were obtained for 171 isolates, including 158 isolates originally identified as *Coemansia* and six Dimargaritales, with 18 of the 171 being type cultures. Assembly quality was variable (FIG 2, SUPP TABLE 1), with the number of contigs ranging from 273 to 10,284, BUSCO completeness ranging from 38.8% to 91.8% (median 73.2), and average coverage ranging from 17.9x to 827.5x (an outlier), median 181x. Genome assembly length was relatively consistent for the Kickxellales species, with most in the expected ~15-25 Mb range. The assembly size for the Dimargaritales species was larger, up to 50.78 Mb (*Tieghemiomyces parasiticus*), even after VizBin filtering. Lower quality (i.e., more fragmented) and higher quality assemblies were more-or-less randomly distributed across the tree (SUPP FIG 2). Correlation plots between the genome assembly variables showed that several variables were significantly correlated (FIG 3). The number of predicted genes was

variable across isolates but was significantly positively correlated with the number of contigs and BUSCO fragmented and duplicated scores. The BUSCO duplicate score was also positively correlated with the BUSCO complete scores. Average coverage was positively correlated with both the BUSCO complete scores and the number of PHYling markers recovered but was negatively correlated with the number of contigs in the assembly and the number of predicted genes. The BUSCO fragmented scores were negatively correlated with the number of PHYling markers recovered.

Phylogenomic reconstruction—We used two different reconstruction methods including FastTree approximately maximum likelihood (ML) (Price et al. 2010) and ASTRAL coalescent-based analyses (Sayyari & Mirarab 2016; Zhang et al. 2018). In addition, we used OrthoFinder (Emms & Kelly 2019) to generate an alternate set of orthologs that was analyzed with FastTree. Single gene alignments and gene trees from the BUSCO ortholog set were evaluated using PHYkit (Steenwyk et al. 2021). Correlation plots showing relationships between alignment and gene tree variables are shown in SUPP FIG 3. The mean bipartition support of the gene trees was most strongly positively correlated with alignment length, followed by the percent of variable sites and the percent of parsimony informative sites. Most alignments were 1500 bp long or less and included >100 taxa. Alignments with more spurious taxa (defined as taxa on branches that are ≥ 20 times the median length of all branches, Shen et al. 2018) generally included fewer total taxa than those with fewer spurious taxa. Likewise, trees with fewer spurious taxa had higher mean bipartition support. Greater numbers of spurious taxa were also associated with lower percentages of variable and parsimony

informative sites and lower treeness scores. The majority of the dataset had no spurious taxa (SUPP FIG 3).

The FastTree phylogeny based on the concatenated amino acid alignment of 511 BUSCO marker genes and 193 taxa is shown in SUPP FIG 4. FIG 5 shows the ASTRAL coalescent-based species tree made from the best ML trees from 461 of the 511 gene trees. There were 50 fewer input gene trees than alignments due to the poor quality of some gene trees (e.g., too few taxa, no resolution among clades). The same four main clades of Kickxellaceae found in the nine-gene phylogeny were also recovered with this larger dataset using both the concatenated ML and coalescent-based analyses. The topologies of the ASTRAL and concatenated ML trees were similar overall to the nine-gene phylogeny (FIG 1) except that Spiromycetales were placed as sister to the “other Kickxellales” + Kickxellaceae clades (rather than sister to Harpellales). All branches received full support in both the ASTRAL and concatenated analyses, except where noted (FIG 4, SUPP FIG 4). ASTRAL quartet scores for the deep nodes of the tree indicated that a large proportion of gene trees had alternative topologies (FIG 4). Nonetheless, in all cases the main topology received ≥98 local posterior support.

The Orthofinder analyses found that 98.3 percent of genes (out of 1,370,791 total for all isolates) were placed among 21,993 orthogroups. One percent of genes were placed in species-specific orthogroups while 1.7 percent of genes were unplaced. Seventy-eight orthogroups had representative genes from 193 taxa and the concatenated alignment was comprised of 101 orthogroups. Unlike the ASTRAL and concatenated ML trees, the OrthoFinder analyses (FIG 5) recovered the same

placement of Spiromycetales as the nine-gene phylogeny (sister to Harpellales). However, the remaining topology of the OrthoFinder tree was the same as the other trees. Venn diagrams depicting the overlap in orthogroups that were detected in each clade found 5,263 (50.4%) orthogroups in common among the Kickxellaceae clades (SUPP FIG 5 A) but only 3,648 (18.5%) orthgroups in common among all of the Kickxellomycotina taxa (SUPP FIG 5 B). The Kickxellaceae clades had between 883 (8.5 %) and 1,329 (12.7%) clade-specific orthogroups (SUPP FIG 5 A), whereas the other orders of Kickxellomycotina had between 805 (4.1%) and 6,173 (31.3%) clade-specific orthogroups (SUPP FIG 5 B). The rarefaction plot of sampled orthogroups by clade (SUPP FIG 6) showed that sampling for the *Coemansia* clade was the most complete, followed by the Kickxellales2 clade. The remaining clades were undersampled in comparison.

Genome annotation—Functional annotation of secondary metabolites (SM), select proteases, and CAZymes were analyzed for each major clade (FIG 6). All four of the Kickxellaceae clades (i.e. *Coemansia*, Kickxellales1, Kickxellales2, and *Kickxella*) had similar types and numbers of SM (FIG 6 A). Each of these four clades had low numbers of NRPS, NRPS-like, siderophores, terpenes, and beta-lactone-containing protease inhibitors. The number of identified subtilases and fungalysins (0 - 9) was likewise consistent across these four clades. The “other Kickxellales” also had few SM identified but varied widely in the number of proteases and subtilases. As expected based on the one previously analyzed genome sequence (Ahrendt et al. 2018), the mycoparasitic Dimargaritales had much higher numbers of NRPS and NRPS-like SM, as well as higher numbers of subtilases. However, *Ramicandelaber brevisporus* (putative

saprotroph) had the highest number of subtilases of any isolate. The number of fungalysins detected in the mycoparasitic Dimargaritales genomes was not much greater than for the “other Kickxellales” (FIG 6 A). The arthropod associated Harpellales had low numbers of all SM and fungalysins, but relatively high numbers of subtilases. The most fungalysins (n=19) were detected from *Dipsacomyces acuminosporus* (putative saprotroph, “other Kickxellales”), followed by *Tieghemiomyces parasiticus* (n=16) and *Dimargaris xerosporica* (n=12) (mycoparasites, Dimargaritales). *Tieghemiomyces parasiticus* had the most subtilases (n=40) followed by *D. cristalligena* RSA 1219 (n=28) (Dimargaritales). No beta-lactone containing protease inhibitors were predicted for any isolates outside of the four Kickxellaceae clades.

The CAZyme profiles were also similar across the four Kickxellaceae clades (FIG 6 B). Although the total numbers differed due to sampling intensity and varying assembly quality within each clade, the relative proportion of each CAZyme family was consistent. Of the families containing known cellulases and pectinases, only glycoside hydrolase (GH) family GH5 and GH3 were detected. However, GH5 was one of the most abundant families among the Kickxellaceae clades. None of the major pectinase-containing gene families (GH28, GH53, GH93), polysaccharide lyases (PL) (PL1, PL3, PL4, PL11), or carbohydrate esterases (CE) (CE8 and CE13) were detected from the Kickxellaceae clades. Kickxellaceae isolates had unexpectedly high numbers of chitinases and laccase-like oxidases, but completely lacked any PL (SUPP TABLE 1). In the entire low coverage dataset, only *Linderina, Mycoëmilia*, and *Spiromyces* had any PL identified. The Dimargaritales had lower numbers of CAZymes than Kickxellales

across all families of enzymes except for carbohydrate binding module 18 and auxiliary activity family 11, both of which have known chitinase functions (Hartl et al. 2012).

We plotted the CAZyme:protease ratio of each major lineage (FIG 7) to compare results between lineages and also compare our results to those obtained by Ahrendt et al. (2018) who found parasitic species generally had a higher ratio of proteases to CAZymes. The linear model tests found significant relationships between these variables for all clades except “other Kickxellales”, a nonmonophyletic group for which there were only 5 representatives. Furthermore, the R^2 values were 0.6267 or greater for three of the Kickxellaceae clades, Harpellales, and Dimargaritales, but the correlation was lower for the *Coemansia* clade (R^2 =0.3663).

DISCUSSION

Low coverage genome sequencing (LCGS) data for phylogenomics and comparative analyses—We used LCGS methods to generate data for 171 Kickxellomycotina isolates, including six mycoparasitic species of Dimargaritales grown from dual cultures that included both the parasites and their host fungi. Our assembly statistics (FIGS 2, 3) and alignment statistics from the BUSCO marker genes (SUPP FIG 3) reveal several interesting features. First, there was a wide variation in quality of both the genome assemblies and the gene alignments. Although average coverage was significantly negatively correlated with the number of contigs and BUSCO fragmented scores, the correlation (~ -0.23 for both) was weaker than anticipated (FIG 3). Zhang et al. (2018) also found a positive correlation between sequencing coverage and BUSCO complete scores, but with high variation among taxa depending on genome size. These findings demonstrate that even for genomes that are small and in a similar size range

(~15-30 Mb), higher coverage does not automatically result in a high-quality assembly. Despite having a median coverage of 181x, our BUSCO complete scores had a median of 73.2. One possible explanation for this discrepancy is amplification and sequencing bias resulting in unequal representation across the genome, which is significantly affected by G:C ratio and can vary based on the specific kits and methods used (e.g., Modlin et al. 2021, Rhodes et al. 2014, Sato et al. 2019). Future work evaluating these biases for fungal genomes specifically will help improve methodology and interpret patterns of coverage versus assembly.

Second, the number of predicted genes was significantly positively correlated with the BUSCO duplicate scores, BUSCO fragmented scores, and contig count, but negatively correlated with average coverage and BUSCO complete scores (FIG 3). This suggests that some of the more fragmented assemblies likely had more erroneous gene predictions. However, several of the isolates with the highest BUSCO duplicate scores were species of Dimargaritales that also had relatively low fragmented scores (*Dispira simplex*: 59.4 % dup., 2.4% frag., *Tieghemiomyces parasiticus*: 66.4% dup., 2.4% frag.). *Dispira simplex* and *T. parasiticus* were the only two isolates with >50% duplicate scores. The one previously published Dimargaritales genome (*D. cristalligena* RSA 468) was suggested to be non-haploid (Ahrendt et al. 2018), therefore it seems probable that these high duplicate scores for the Dimargaritales may indicate a non-haploid state for these taxa. Although there are still relatively few reports of whole genome duplications across Fungi (Albertin & Marullo 2012), there is evidence for this phenomenon among Mucoromycota fungi (Corrochano et al. 2016; Ma et al. 2009). Wang et al. (2018) also suggested genome duplication as a possible explanation for the much larger genome

sizes of Harpellales gut fungi compared to other Kickxellomycotina. Recently, estimates of heterozygosity based on genome sequence data suggested that several Zoopagomycota species were at least diploid, including *Coemansia reversa, Linderina pennispora, Martensiomyces pterosporus*, and *Ramicandelaber brevisporus* (Amses et al. 2022).

Finally, although only 36.3% of our assemblies had BUSCO complete scores ≥80% (SUPP FIG 2), we were still able to retrieve hundreds of marker genes from even the poorest assemblies for phylogenomic reconstruction (FIG 2). Quality among the 511 marker gene alignments was also variable (SUPP FIG 3), but the different phylogenetic reconstruction methods nonetheless largely resulted in the same topologies for the various phylogenies we reconstructed (FIG 4, SUPP FIG 4). The LCGS data also worked well as input for OrthoFinder analyses, which provided a different set of loci for reconstruction (FIG 5). Interestingly, the OrthoFinder analysis was the only method that recovered the same relationships between “other Kickxellales” and placement for Spiromycetales (sister to Harpellales) as the nine-gene dataset (which had a more complete taxonomic representation), supporting the idea that OrthoFinder analyses may be more robust to missing taxa in addition to missing genes (Emms & Kelly 2015).

Phylogenetic Relationships among Kickxellomycotina—Our nine-gene phylogeny reconstructed some unexpected relationships among the trichomycetes (Asellariales, Barbatosporales, Harpellales, Orphellales) and the Spiromycetales. Contrary to the eight-gene analyses by Tretter et al. (2014) which resolved the insect-symbiotic Kickxellomycotina in several clades, our analyses recovered a monophyletic clade of insect gut fungi, with Asellariales, Harpellales, Orphellales, and Barbatosporales

resolved as one evolutionary unit (FIG 1). Spiromycetales are nested within the gut fungi clade in our analyses, rather than as a separate clade as found previously (Tretter et al. 2014; White et al. 2006b). Species of *Spiromyces* and *Mycoëmilia scoparia* are considered saprotrophic due to their growth in axenic culture, but their trophic modes require further empirical testing. *Mycoëmilia scoparia* was originally isolated from soil containing dead isopods and grew vigorously on several different nutrient media (Kurihara et al. 2004). This rapid growth in axenic culture suggests that *M. scoparia* is a saprobe, but there is wide variation in the saprobic ability of insect-symbiotic fungi in Kickxellomycotina. For example, some Harpellales gut fungi that are considered obligate symbionts of insects can grow well in axenic culture (i.e., *Smittium* spp., Wang et al. 2017) whereas Asellariales and Orphellales gut fungi that inhabit isopod and stonefly nymph hosts have not been successfully cultured (Valle & Cafaro 2008; White et al. 2018). It is possible that *M. scoparia* is an intermediate form that is a facultative symbiont of isopods with saprotrophic capabilities. A recently described genus and species, *Unguispora rhaphidophoridarum,* matches this intermediate form which the authors term an “amphibious” life cycle (Ri et al. 2022). *Unguispora* adheres to the proventriculus of cave crickets and produces secondary spores shed into the cricket gut which subsequently grow saprophytically on cricket dung (Ri et al. 2022). However, *U. rhaphidophoridarum* was placed within the Kickxellaes (sister to *Linderina*) rather than the Spiromycetales in the molecular phylogeny. On the other hand, *Spiromyces* spp. have only been isolated from rodent dung, and they exhibit fastidious growth in culture but improved growth in mixed cultures containing bacteria and other dung-associated organisms (Benjamin 1963; O’Donnell et al. 1998). Interestingly, *S. aspiralis* had a

higher protease:CAZyme ratio compared to *M. scoparia* (FIG 6), which is a profile more similar to biotrophs, but additional data are needed to interpret this finding.

If our topology more accurately reflects the relationships among these orders of arthropod-associated and putative saprotrophic fungi, then it implies a different evolutionary scenario than what was implied by Tretter et al. (2014). The topology recovered by Tretter et al. (2014) suggested that the arthropod gut fungi likely evolved independently at least three times within Kickxellomycotina. In contrast, our topology suggests alternative hypotheses: 1) a single origin of biotrophic gut fungi with a putative reversion to saprotrophy (Spiromycetales), or 2) two independent origins of gut fungi with the Spiromycetales retaining the ancestral saprotrophic state. Either of these scenarios would be more parsimonious than three independent origins of biotrophy with arthropods. However, given the poor taxon sampling of taxa in the “other Kickxellales" clade (SUPP FIG 6) and a lack of sequence data for many arthropod gut fungi, inference about the evolutionary origins of these fungi will benefit from additional genome sequencing. Future sequencing efforts paired with ancestral state reconstructions will help to resolve these questions.

One of the other main hypotheses we aimed to test was regarding the relationship between Dimargaritales, Ramicandelaberales, and the Kickxealles. Previous analyses suggested that either Dimargaritales and Ramicandelaberales were sister taxa (Tretter et al. 2014) or that Dimargaritales were the earliest diverging lineage in Kickxellomycotina and that Ramicandelaberales was sister to the remaining taxa (Davis et al. 2019). However, these previous phylogenomic analyses (e.g., Ahrendt et al. 2018; Chang et al. 2015; Davis et al. 2019; Spatafora et al. 2016) included sampling

from across Kingdom Fungi and had limited taxon sampling from Kickxellomycotina. In an eight-gene analysis focused on Kickxellomycotina, Tretter et al. (2014) had greater taxon sampling, but were still unable to resolve the placement of Dimargaritales and Ramicandelaberales. Additionally, we aimed to test the species complex comprised of several *Coemansia* species found by Chuang et al. (2018). All analyses of our 171 low coverage genomes resolve Dimargaritales as the earliest diverging lineage in the subphylum and Ramicandelaberales as sister to the rest of Kickxellomycotina (FIGS 1, 4, 5, SUPP FIG 4). As far as the species complex, *Coemansia pectinata* (IMI 142377), *C. furcata*, and *C*. "*aciculifera*" from Taiwan were all placed in separate, monophyletic clades (Kickxellales2 clade). Unfortunately, we were unable to obtain successful LGCS for *C. pennisetoides* but the rDNA data included in the nine-gene dataset placed it in a separate clade from the rest.

Among the other saprotrophic species, our analyses revealed that Kickxellaceae isolates are divided among four large, monophyletic clades ("Kickxellaceae clades"). These relationships solidify previous findings that *Kickxella* and *Spirodactylon* are nested within *Coemansia* sensu lato (FIG 1) (Chuang et al. 2018). There were no broad phylogenetic patterns associated with either morphology, geography, or substrate across these clades. There is a possibility that some species are endemic to certain regions because some clades were comprised entirely of isolates from one country (SUPP TABLE 2), but this cannot be confirmed due to sampling bias. Nonetheless, as with Mortierellaceae, LCGS data provided resolution where morphology and rDNA data could not (Vandepol et al. 2020). We have begun the process of reorganizing the taxonomy of Kickxellaceae to reflect these phylogenetic relationships by emending the

genera *Coemansia* and *Kickxella* and transferring the type of *Spirodactylon* (*S. aureum*) to *Coemansia* (i.e. synonymizing *Spirodactylon* to *Coemansia*) (SUPP INFO, SUPP TABLE 2). However, much work remains to be done, including creating new genera for the Kickxellales1 and Kickxellales2 clades, new species descriptions for the many unidentified Kickxellaceae, and reorganization of Dimargaritales genera to recognize monophyletic relationships.

Finally, the orthogroup rarefaction plot (SUPP FIG 6) indicates that all the clades except *Coemansia* are undersampled. Our sampling includes representatives of all putative saprotrophic genera of Kickxellales (except *Myconymphaea*). Therefore, the incomplete sampling indicated in the rarefaction plot suggests that there is likely significant undiscovered diversity in these groups. For example, several of the genera within the "other Kickxellales" clade are monotypic and/or are rarely reported in the literature. The apparent rarity of these taxa may be tied to their geographic distributions; many of these fungi are from tropical sites that remain highly undersampled. The monotypic *Dipsacomyces acuminosporus* was collected from Honduras (Benjamin 1961), *Linderina* spp. have been collected from China, India, and Liberia (Chang 1967; Raper & Fennell 1952), and the monotypic *Martensiomyces pterosporus* is from the Democratic Republic of the Congo (Meyer 1957) (SUPP TABLE 2). Any future microfungal biodiversity surveys in these areas will certainly uncover many new Kickxellomycotina species and lineages.

Functional annotations of genomes from across the Kickxellomycotina phylogeny—Previous reports proposed that some species of Kickxellales may not be saprotrophic. Mycoparasitic interactions have been suggested for *Coemansia reversa,*

which was found growing on *Isaria* spp. (Ascomycota) (Bainier 1906; Linder 1943) after its original description from rat dung (van Tieghem & Le Monnier, 1873). *Martensella corticii* is considered a mycoparasite of *Vesiculomyces citrinus* (=*Corticium radiosum*) (Basidiomycota) (Linder 1943) and was consistently found only on this host during a large survey of *Corticium* species (Jackson & Dearden 1948). Although no molecular data or cultures exist for *M. corticii*, its morphology unambiguously places it within the Kickxellales rather than Dimargaritales due to the formation of multicelled sporocladia with pseudophialides (Linder 1943). Conversely, potential arthropod associations have been previously suggested for other Kickxellales species. For example, *Linderina macrospora* (NBRC 105416 = BTCC-F30) and *C. javaensis* (NBRC 105414 = BTCC-F33) grew and sporulated better on media supplemented with aphids (Kurihara et al. 2008) and *Pinnaticoemansia coronantispora* may inhabit the foregut of earwigs (Dermaptera) (unpublished data, Y. Degawa pers. comm. 2021). All of these observations led us to investigate differences in the functional annotations of our Kickxellomycotina genomes to better understand the likely trophic modes among the clades.

Proteases and secondary metabolites

Fungalysins (MEROPS family M36) are metalloproteases that have been implicated in pathogenesis of opportunistic human pathogens (e.g., *Aspergillus fumigatus, Cryptococcus neoformans*) and are suggested to degrade extracellular matrix proteins like elastin and keratin (Brouta et al. 2002; Markaryan et al. 1994; Pombejra et al. 2018). *Batrachochytrium dendrobatidis,* the chytrid pathogen of amphibians, has a large expansion of the M36 gene family that were differentially

expressed during different stages of the life cycle (Rosenblum et al. 2008). In plant pathogenic species (e.g., *Colletotrichum graminicola*), fungalysins function to cleave plant chitinases and may act as effectors (Ökmen et al. 2018; Sanz-Martín et al. 2015). An expansion of genes belonging to the M36 family was also detected in the non-pathogenic *Coprinopsis cinerea* (Basidiomycota) (Lilly et al. 2008) although the function in this saprotroph is unclear. In our dataset, the mycoparasitic Dimargaritales had more fungalysins than other Kickxellomycotina, which is consistent with the findings that the M36 family is often associated with pathogens. However, the species with the most predicted fungalysins (n=19) was *Dipsacomyces acuminosporus*, a putative saprotroph. The *D. acuminosporus* assembly had a BUSCO complete score of 77.8, a duplicate score of 0.5, and a 6.1 fragmented score, so the assembly was of moderate quality. Due to the correlation between BUSCO fragmented scores and number of predicted genes (FIG 3), some of these predicted fungalysins may be erroneous. On the other hand, little is known about *D. acuminosporus* because it has only been isolated once from soil (Benjamin 1961; Young 1999), highlighting the need for further studies of the poorly understood ecology of many Kickxellomycotina.

The subtilase family (Pfam PF00082) is diverse and is the second largest serine protease family (Siezen & Leunissen 1994). Within fungi these enzymes have been associated with many different roles. Some of the best characterized functions have been described for fungal pathogens and parasites (e.g., mycoparasitic and nematode-parasitic Hypocreales, Ascomycota) (Iqbal et al. 2018). Results from large-scale comparative genomics analyses have led to the hypothesis that subtilase gene family expansions are adaptations to utilize animal tissues as a food source, at least among

Ascomycota (Muszewska et al. 2011). For example, the entomopathogenic fungus *Metarhizium robertsii* (Hypocreales) had 48 subtilase genes, which was the most out of 83 analyzed genomes (Li et al. 2017). A comparison of Zoopagomycota subtilases also found expansions among mycoparasitic species (Ahrendt et al. 2018), a result that was recapitulated with our data (FIG 6 A). Kickxellales taxa only had a small number of predicted subtilases, but on average had more predicted subtilases than fungalysins. Species of Harpellales had the second highest average number of subtilases. The Harpellales are associated with insects but are considered commensal rather than pathogenic (except in the case of *Smittium morbosum*, Wang et al. 2017), and the putative role of subtilases in these associations remains unknown.

Fungi are also important producers of secondary metabolites (SM) that perform a variety of functions (e.g., toxins), some of which have application in pharmaceuticals (e.g., antibiotics, antifungals, statins) (Cox & Simpson 2009; Rokas et al. 2018). It has also been suggested that SM can function defensively to deter fungal grazers, such as invertebrates (Kempken and Rohlfs 2009). Most groups of SM are non-ribosomal peptide synthetases (NRPS and NRPS-like), polyketide synthases (PKS and PKS-like), or terpene cyclases (Macheleidt et al. 2016). Before many genome sequences of early diverging fungi (i.e., Blastocladiomycota, Chytridiomycota, Microsporidia, Mucoromycota, Rozellomycota, Zoopagomycota) were available, little was known about SM in those groups, but they were generally thought to be deficient in SM (Bushley & Turgeon 2010). However, recent analyses of Zoopagomycota fungi revealed that *Dimargaris cristalligena* RSA 468 (Dimargaritales) had the second highest proportion of predicted SM in Zoopagomycota (after *Basidiobolus meristosporus*), followed by

Linderina pennispora ATCC 12442 and *Martensiomyces pterosporus* CBS 209.56 (Kickxellales) (Tabima et al. 2020). We found support for the hypothesis that *Coemansia* and *Kickxella* isolates would have relatively few predicted SM, with increased numbers in other Kickxellales, and the most SM among the Dimargaritales (Ahrendt et al. 2018; Tabima et al. 2020) (FIG 6 A). We found that SM predictions were consistent across the Kickxellaceae and "other Kickxellales" clades with only one or two NRPS and NRPS-like clusters identified in all taxa. Almost none of the predicted SM from any isolate had similarity to known clusters (i.e., the antiSMASH "KnownClusterBlast" found no matches). As expected, Dimargaritales had far more predicted SM than all other Kickxellomycotina. *Tieghemiomyces parasiticus* had the most NRPS and NRPS-like proteins (n=89) whereas *Dispira simplex* had the least (n=25). In comparison, species of *Aspergillus* (characterized as a SM-rich genus) were predicted to have between 39 and 81 SM clusters (Inglis et al. 2013). Kickxellalaceae clades consistently had 1–3 terpenes and 1–5 siderophores. Forty-nine Kickxellalaceae isolates had one beta-lactone containing protease inhibitor detected. No beta-lactone containing protease inhibitors were predicted for any Kickxellomycotina in other clades, so these proteins may be an adaptation specific to Kickxellaceae. Many beta-lactone compounds have characterized antimicrobial and antitumor properties and they can target a wide array of substrates (Robinson et al. 2019), so these enzymes are of particular interest in pharmacological studies.

Carbohydrate active enzymes (CAZymes)

To gain a better understanding of the substrates Kickxellales fungi might be able to utilize for nutrition and growth, we analyzed the predicted CAZymes present in each

major clade of Kickxellomycotina (FIG 6 B). CAZymes are categorized into carbohydrate esterases (CE), carbohydrate binding modules (CBM), glycoside hydrolases (GH), glycosyl transferase (GT), polysaccharide lyases (PL), and auxiliary activities (AA). Similar to the results for SM and proteases, the relative proportion of predicted CAZyme families was fairly constant across taxa in Kickxellaceae clades, although the total numbers differed. We were specifically interested in pectinases, cellulases, and chitinases. While previous analyses have found a complete loss of the pectinase-containing family GH28 in *Coemansia reversa* NRRL 1564 (Chang et al. 2015) our analyses extended this finding and suggest a complete loss of GH28 across the entire subphylum (SUPP TABLE 1). Indeed, none of the major pectinase families (GH53, GH93, PL1, PL3, PL4, PL11, CE8, CE13) were detected among Kickxellomycotina genomes and no PL families of any kind were found from any isolates except *Linderina* spp., *Mycoëmilia scoparia*, and *Spiromyces aspiralis*. Similarly, out of the main cellulase families (AA9, GH5, GH6, GH7, GH12), only AA9 and GH5 were predicted across all the isolates. However, enzymes in the GH5 family are diverse and facilitate the degradation of other substrates, such as starches and other polysaccharides (Davies & Attia 2021). β-glucosidases from the GH1 and GH3 families are known to degrade cellobiose in solution (Sørensen et al. 2013; Payne et al. 2015) and the GH3 family was detected in nearly all isolates.

Chitinases are important for endogenous cell wall remodeling as well as utilization of exogenous chitin compounds (Seidl 2008; Hartl et al. 2012). However, chitinases are thought to play a role in mycoparasitism, for example among *Trichoderma* species (Ascomycota) (Benítez et al. 2014). Chitinases are known from

families CE4, CE9, GH18, GH19, GH20, GH46, GH75, GH80, and AA11 (Battaglia et al. 2011; Latgé 2007; Seidl 2008). Ahrendt et al. (2018) reported family GH19 from Zoopagomycota for the first time and found that the AA11 family was only detected among mycoparasites. Furthermore, *D. cristalligena* RSA 468 was unique in having several CBM18 genes. Once again, CBM18 was found almost exclusively among the mycoparasitic Dimargaritales, but *Spiromyces aspiralis* RSA 2271 also had one predicted CBM18 gene. Most isolates had several GH19 and CE9 and also numerous GH46 genes predicted, but GH75 and GH80 genes were absent. However, unlike the results from Ahrendt et al. (2018), we found an expansion of AA11 across all Kickxellomycotina, except in Harpellales which had few predicted AA11 genes (FIG 6 B). Chitinases had not been previously characterized for Kickxellales fungi, but we found diverse chitinase enzymes in numbers similar to those found in *Rhizopus oryzae* (Mucoromycota) (Battaglia et al. 2011). *Rhizopus oryzae* is a generalist saprotroph capable of utilizing a wide variety of substrates. Finally, we also unexpectedly found numerous AA1 genes predicted in the Kickxellaceae clades. This family of enzymes contains known fungal laccases as well as ferroxidases and laccase-like multicopper oxidases. Fungal laccases are best characterized among Dikarya and can degrade a wide variety of compounds including phenolics (Baldrian 2005). The most well-known laccases are those from Basidiomycota white and brown rot fungi which break down lignin (families AA1_1 and AA2) and hemicellulose (Hage & Rosso 2021). The predictions for Kickxellomycotina did not specify which subfamily the AA1 genes might belong to, but the detection of AA2 (fungal class II peroxidases) was surprising because lignin degradation among early diverging fungi has been poorly characterized and has

been considered unlikely (Floudas et al. 2012; Janusz et al. 2017). However, some reports have shown activity and/or presence of these and other ligninolytic enzymes in Mucoromycota (e.g., *R. oryzae* – Freitas et al. 2009, *Mucor racemosus* – Bonugli-Santos et al. 2010) and Chytridiomycota (Lange et al. 2019) but were missing from mycorrhizal Endogonaceae (Mucoromycota) genomes (Chang et al. 2019). Unfortunately, the specific enzyme families involved with lignin decay have not been well characterized in most early diverging fungal species. Interestingly, this is the first report of these putative lignin-degrading enzymes among any Zoopagomycota taxa and therefore warrants further investigation.

The patterns that we detected in the CAZyme:protease ratios (FIG 7) were consistent with the known trophic ecology of Kickxellomycotina. The ratios for the Kickxellales more closely match saprotrophic profiles (with greater CAZymes than proteases), while Dimargaritales and Harpellales have the opposite profile which aligns with other parasitic and animal-associated species (Ahrendt et al. 2018). Taken together, these results suggest that Kickxellales (particularly taxa in Kickxellaceae) have generalist saprotrophic capabilities, although the potential for facultative interactions with other fungi or insects cannot be ruled out. Species in Kickxellaceae commonly grow on dung, and the predicted enzymatic profiles fit the utilization of both plant and potentially fungal or insect substrates that are often present in the diets of herbivorous and omnivorous animals. Another interesting niche is within fungal necromass communities, which play an important role in C and N cycling while degrading mycelium (Zhang et al. 2018b). For example, *Ramicandelaber* was detected among necromass communities in Minnesota and was found to be significantly impacted by necromass

substrate quality (Beidler et al. 2020). Whether other Kickxellomycotina are important in fungal decomposition remains to be investigated.

Material and Methods

Cultures—Lyophilized cultures were obtained from the University of Florida (R.K. Benjamin (RSA) and G.L. Benny collections), United States Department of Agriculture Research Service Culture Collection (NRRL), the Westerdijk Fungal Biodiversity Institute (CBS), Bioresource Collection and Research Center, Taiwan (BCRC), National Institute of Technology and Evaluation Biological Resource Center, Japan (NBRC), and the Centre for Agriculture and Biosciences International, UK (IMI) culture collections (SUPP TABLE 2). We refer to individual cultures as isolates (e.g., RSA 532 is an isolate of *Coemansia*). Lyophililzed cultures previously preserved in sealed glass vials were scored with a file, broken open, and dried spores and hyphae of each fungal isolate was placed into ~20 mL sterile MEYE broth (3 g malt extract, 3 g yeast extract, 5 g peptone, 10 g dextrose, and distilled water (dH_2O) up to 1 L). Once hyphal growth was observed in the broth, tissues were transferred to either MEYE agar (as above but with 18 g agar), V8 agar (163 mL V8 juice, 3 g $CaCO_3$, 18 g agar, and dH_2O up to 1 L), or YGCH agar (10 g yeast extract, 15 mL glycerol, 15 g casein hydrolysate, 1.0 g K_2HPO_4, 0.5 g $MgSO_4 \cdot 7H_2O$, 18 g agar, and dH_2O up to 1 L) plates supplemented with mixtures of different antibiotics (Benny et al. 2016). Various antibiotic combinations were used in an attempt to clear bacterial contamination found in a few cultures. Multiple replicate plates were grown for each isolate in either an 18°C incubator or at room temperature, depending on the optimal growth conditions for each isolate. Once each isolate was well established on an agar plate, sporulating hyphae were scraped from the surface and

placed in 2x CTAB (cetyltrimethylammonium bromide) buffer for DNA extraction. Most isolates required several replicate plates to obtain sufficient material for DNA extraction. Species identification was based on the identification made by the collector at the time of isolation.

DNA and RNA extraction and Genome sequencing—DNA extraction followed the CTAB protocol of Gardes and Bruns (1993), but with the following modifications: tissues were ground using a micropestle attached to a drill press and subjected to two rounds of freezing and thawing, the phenol:chloroform wash step was followed by an additional chloroform wash step, and samples were left at -20°C overnight for the isopropanol precipitation step. After extraction, DNA was treated with RNase A and quantified by both Nanodrop 2000 spectrophotometer (ThermoFisher Scientific, Waltham, MA) and Qubit 4.0 Fluorometer (Invitrogen, Waltham, MA). The large subunit rDNA (28S) of several representative *Coemansia* isolates was amplified with primers LROR/LR5 (Hopple & Vilgalys 1994; Vilgalys & Hester 1990) and sent for Sanger sequencing at GeneWiz (South Plainfield, NJ). RNA extraction for *Pinnaticoemansia coronantispora* was performed using a Zymo Direct-zol RNA kit (Irvine, CA) and TRI Reagent (Sigma-Aldrich, St. Louis, MO). RNA was treated with DNase I, quantified with Qubit using the RNA kit, and stored in a -80°C freezer until shipment to the Joint Genome Institute (JGI).

Most genomic DNA samples were sent to the JGI for either low coverage Illumina (San Diego, CA) sequencing or full coverage reference genome sequencing with PacBio (Pacific Biosciences, Menlo Park, CA) (see details below). A set of 25 samples were sent to Novogene (Sacramento, CA) for low coverage Illumina sequencing (SUPP

TABLE 1). At Novogene, sequencing libraries were generated using NEBNext DNA Library Prep Kit (New England BioLabs, Ipswich, MA) following manufacturer's recommendations using 1.0 µg DNA per sample and indices added to each sample. Genomic DNA was randomly fragmented to a size of 350 bp by shearing, then DNA fragments were end polished, A-tailed, and ligated with the NEBNext adapter for Illumina sequencing, and further PCR-enriched by P5 and indexed P7 oligos. The PCR products were purified (AMPure XP system) and resulting libraries were analyzed for size distribution by Agilent 2100 Bioanalyzer and quantified using real-time PCR. Sequencing was performed on the Illumina NovaSeq 6000 sequencing platform with NovaSeq XP v1 reagent kits, following 2x150 indexed chemistry.

Library preparation for Illumina sequencing at the JGI utilized plate-based DNA preparation performed on the PerkinElmer (Waltham, MA) Sciclone NGS robotic liquid handling system using Kapa Biosystems (Wilmington, MA) library preparation kit. A Covaris (Woburn, MA) LE220 focused-ultrasonicator sheared 200 ng of sample DNA to 600 bp. Sheared DNA fragments were size-selected by double-SPRI (solid-phase reversible immobilization beads) and then the selected fragments were end-repaired, A-tailed, and ligated with Illumina compatible sequencing adapters from Integrated DNA Technologies (Coralville, IA) containing a unique molecular index barcode for each sample library. Libraries were quantified using Kapa Biosystem’s next-generation sequencing library qPCR kit and run on a Roche Diagnostics (Indianapolis, IN) LightCycler 480 real-time PCR instrument. Quantified libraries were then multiplexed with other libraries, and the pool of libraries was prepared for sequencing on the

Illumina NovaSeq 6000 sequencing platform using NovaSeq XP v1 reagent kits, S4 flow cell, following a 2x150 indexed paired end run recipe.

The *Pinnaticoemansia coronantispora* CBS 131509 genome was prepared for PacBio sequencing using >10kb Blue Pippin Size Selection, unsheared preparation. Using SMRTbell Template Prep Kit 1.0 (Pacific Biosciences), 1500ng of genomic DNA was directly treated first with exonuclease to remove single-stranded ends and then with DNA damage repair mix followed by end repair and ligation of blunt adapters. The final library was size-selected with BluePippin system (Sage Science, Beverly, MA) with a 6 kb cutoff size and purified with AMPure PB beads. PacBio Sequencing primers were then annealed to the SMRTbell template library and sequencing polymerase was bound to them using Sequel II Binding kit 2.0. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosystems' Sequel II sequencer using v4 sequencing primer, 8M v1 SMRT cells, and Version 2.0 sequencing chemistry with 1x1800 min sequencing movie run times. Filtered PacBio CCS reads were then assembled with Flye version 2.7.1-b1590 (https://github.com/fenderglass/Flye) to generate an assembly and polished with gcpp --algorithm arrow version SMRTLINK v8.0.0.80529 (https://www.pacb.com/support/software-downloads). The genome was annotated using the JGI Annotation pipeline (Grigoriev et al. 2014).

For the transcriptome, plate-based RNA sample prep was performed on the PerkinElmer Sciclone NGS robotic liquid handling system using Illumina's TruSeq Stranded mRNA HT sample prep kit utilizing poly-A selection of mRNA following the protocol outlined by Illumina (https://support.illumina.com/sequencing/sequencing_kits/truseq-stranded-mrna.html).

The total RNA starting material was 1 ug per sample and 8 PCR cycles were run for library amplification. The prepared library was then quantified and sequenced with Illumina NovaSeq 6000 as above. Filtered RNA-Seq reads were assembled into consensus sequences using Trinity v.2.11.0 (Grabherr et al. 2011).

Low coverage genome assembly and annotation—Sequence data were assembled using the Automatic Assembly for the Fungi (AAFTF) pipeline (Stajich et al. 2021). This pipeline trims reads using trimmomatic v0.33 (Bolger et al. 2014), filters contaminating reads (e.g., PhiX) with BBmap v38.16 (Bushnell 2014) and performs de novo assembly with SPAdes v3.9.0 (Bankevich et al. 2012). The assembly is then further cleaned by removing additional contaminating sequences with sourmash (Brown & Irber 2016) and duplicate contigs with minimap2 (Li 2018), followed by polishing with Pilon (Walker et al. 2014). Assemblies were then annotated using funannotate v1.8.1 (Palmer & Stajich 2020). Funannotate first masks repetitive regions of the genome with RepeatMasker (Smit et al. 2015) and creates *ab initio* gene prediction consensus models with EVidenceModeler (Haas et al. 2008) after training Augustus v3.3 (Stanke et al. 2006) using the BUSCO (Simão et al. 2015) fungi_odb10 data set and GeneMark (Ter-Hovhannisyan et al. 2008). Exon locations were inferred from protein alignments against the SwissProt (The UniProt Consortium 2021) database with BLASTX and exonerate v2.4.0 (Slater & Birney 2005). Functional annotations were added by searching against antiSMASH (Blin et al. 2019), dbCAN (Yin et al. 2012), eggNOG (Huerta-Cepas et al. 2019), MEROPs (Rawlings et al. 2018), and Pfam (Finn et al. 2014) databases with HMMER (Eddy 2009) and DIAMOND (Buchfink et al. 2021). Of particular interest were proteases (fungalysins MER0001400 and subtilases PF00082)

and secondary metabolites that have been associated with parasitism. Additionally, CAZyme (carbohydrate active enzyme) families containing known chitinases, cellulases, and pectinases were evaluated to gain an understanding of the putative trophic mode of Kickxellales fungi. Box plots and bar graphs of these data were made using the R (R core team 2020) package ggplot2 (Wickham 2016) with the cowplot add-on (Wilke 2020). We also plotted the number of CAZymes versus the number of proteases for each clade. The linear model function in R was used to evaluate the significance of the relationships and calculate the R^2 values.

Assembly statistics (e.g., average coverage, contig counts, etc.) were assessed with bbmap and BUSCO. BUSCO outputs scores for the number of complete, missing, fragmented, and duplicate benchmarking genes found in the genome assemblies as a method of estimating completeness and accuracy of the assembly. The predicted number of genes for each genome was based on the annotation outputs. Correlation plots depicting the relationships between these variables were made in R using the GGally extension (Schloerke et al. 2021) of the ggplot2 package.

Filtering contaminant sequences—Genomes from Dimargaritales mycoparasites co-cultured on host fungi were expected to be contaminated with host sequences. In order to identify and remove host sequences, taxonomic identification and coverage for assembled contigs was assessed with BlobTools v1 (Challis et al. 2020). The taxonomic and coverage data from BlobTools and the assembly fasta files were then input to VizBin (Laczny et al. 2015). The graphical interface of VizBin allows the user to select sequences of interest by drawing polygons around clusters of sequences, thereby filtering the dataset. For each mycoparasite genome, polygons were drawn to exclude

clusters of sequences taxonomically identified as host (i.e., Mucoromycota or Ascomycota) (SUPP FIG 1). The resulting filtered fasta files were reanalyzed with funannotate to obtain the corrected annotations and subsequently used in downstream analyses. Likewise, bacterial contamination of a few *Coemansia* isolates was expected due to observed bacterial growth on some agar plates despite the use of several different antibiotics and transfer attempts. Additionally, we were interested in evidence of putative endohyphal symbionts, as have been identified across Mucoromycota species (Chang et al. 2018; Deveau et al. 2018) but have not been confirmed in any Zoopagomycota. To screen the genomes for bacterial sequences, assemblies were processed with the Autometa pipeline (Miller et al. 2019). The first step is to bin contigs according to their Kingdom-level classification using Prodigal gene prediction (Hyatt et al. 2010) and identity searches against the NCBI NR database. The output fasta files containing contigs binned as eukaryotic were then reannotated with Funannotate as above and used in downstream analyses.

Nine-gene dataset—To build a phylogeny including taxa for which genome data were unavailable (species of *Asellaria, Barbatospora*, *Myconymphaea*, Orphellales, *Spirodactylon*), we obtained reference sequences of seven protein-coding genes (actin, β tubulin, EF1α, MCM7, RPB1, RPB2, and TSR1) and the small subunit (18S) and large subunit (28S) ribosomal DNA from NCBI. These markers (except actin) were generated for many Kickxellomycotina taxa by Tretter et al. (2014) and we wanted to build upon this dataset by capturing these loci from our isolates. The automated Target Restricted Assembly Method (aTRAM) (Allen et al. 2015) was first used to retrieve 18S and 28S rDNA sequences from the low coverage genome data. The aTRAM pipeline creates

BLAST-formatted databases for the forward reads of each sample and indexes the corresponding paired reverse reads. A reference sequence is then used as a query against these read databases to retrieve the best matches. This process is repeated iteratively until a *de novo* contig matching the reference sequence is obtained. *Coemansia* and *Kickxella* reference 18S and 28S sequences included those downloaded from GenBank and Sanger sequences generated as part of this study. Resulting contigs were further refined into scaffolds with CAP3 (Huang & Madan 1999) and ragtag (Alonge et al. 2019).

Next, Hidden Markov models (HMM) profiles were built using HMMER v3.3.2 (Eddy 2011; Eddy et al. 2020) for the protein coding sequences of actin, β tubulin, EF1α, MCM7, RPB1, RPB2, and TSR1 based on reference sequences. The PHYling pipeline was then used to retrieve matching sequences from the LGCS data. This pipeline uses hmmsearch and hmmalign to select the best match to the models and assumes that the highest hits are approximate to orthologs. The pipeline then creates sequence alignments for each locus. Output fasta files of recovered loci from each of the low coverage datasets were imported into Mesquite (Maddison & Maddison 2017), aligned with MUSCLE (Edgar 2004) and adjusted by eye, with unaligned regions manually excluded. Maximum Likelihood (ML) inference of the partitioned dataset was performed with RAxML v8 (Stamatakis 2014) using the GTRCAT substitution model for nucleotides, the LG model for amino acids, and the extended majority rule criterion for automatically determining bootstrap replicates. Tree output was visualized using FigTree v1.4.3 (Rambaut 2016) and edited in InkScape v1.0.2 (https://inkscape.org/en/).

Phylogenomic reconstruction— Three separate methods were used for phylogenetic reconstruction: the PHYling pipeline (utilizing FastTree) (Price et al. 2010), OrthoFinder v2.3.8 (Emms & Kelly 2019), and the ASTRAL coalescent-based method (Sayyari & Mirarab 2016; Zhang et al. 2018). The PHYling pipeline searches the assembled genomes for a set of 758 putatively single-copy markers from the BUSCO dataset (fungi_odb10). Both the PHYling and ASTRAL analyses were based on the same set of 511 recovered (out of 758 total) marker genes. The protein fasta file output for each isolate was searched for the best hit to these markers, and the resulting hits were aligned to the original HMM using hmmalign. Alignments were trimmed with ClipKIT (Steenwyk et al. 2020) and gene trees were reconstructed in RAxML. All individual gene trees output from RAxML were then used as input for analysis in ASTRAL with detailed branch support output. The concatenated alignment was analyzed with FastTree approximately-ML using the LG and Gamma20 models (Price et al. 2010). All 511 BUSCO marker gene alignments (and their corresponding RAxML gene trees) identified by PHYling were analyzed with PHYkit (Steenwyk et al. 2021). Alignments were assessed for length, number of taxa, parsimony informative sites, variable sites, and relative composition variability (lower values may indicate lower composition bias - Phillips & Penny 2003). Gene trees were analyzed for spurious taxa (taxa on branches that are ≥ 20 times the median length of all branches - Shen et al. 2018), treeness (higher treeness values are thought to indicate a greater signal to noise ratio - Phillips & Penny 2003), and bipartition support values.

OrthoFinder was used to identify a different set of marker genes: orthologs identified from among the input taxa. The OrthoFinder analysis was run using

DIAMOND for sequence similarity searching and the multiple sequence alignment method for tree reconstruction using MUSCLE and FastTree for reconstruction. The results were analyzed with KinFin v1.0 (Laetsch & Blaxter 2017) to obtain an orthogroup rarefaction plot by clade and Venny v2.1 (Oliveros 2015) was used to create Venn diagrams of orthogroup overlap between clades. SUPPLEMENTAL TABLE 1 lists all isolates, including full coverage reference genomes, that were included in all analyses.

DATA AVAILABILITY

All genome data have been deposited to Genbank (accession numbers in SUPP TABLE 1) and R scripts and other analysis files are available at the Center for Open Science OSF site: DOI 10.17605/OSF.IO/6ZNP7. Cultures not already deposited in collections were submitted to the CBS (Centraal Bureau voor Schimmelcultures, Westerdijk Fungal Biodiversity Institute) collection.

ACKNOWLEDGEMENTS

This work was supported by the National Science Foundation (DEB 1441677 to MES; DEB 1441715 to JES; DEB-1441604 to JWS), the Department of Plant Pathology, the USDA-NIFA (FLA-PLP-005289), and the Institute for Food and Agricultural Sciences at the University of Florida. The work (proposal 10.46936/10.25585/60001062) conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. We would also like to thank undergraduate researchers who helped along the way: Sophia Torla, Adrian Emery, and Elizabeth Dautel.

Figures and Tables

Figure 1. RAxML tree inferred from seven protein coding genes (actin, β tubulin, EF1α, MCM7, RPB1, RPB2, TSR1) and two rDNA (18S and 28S) genes. The LG model was used for amino acids and GTRCAT for nucleotides. Bolded branches had ≥75% bootstrap support. Major clades are outlined with boxes, isolates represented by low coverage genome data are colored blue and data downloaded from the Joint Genome Institute or GenBank are in black. The red arrows highlight the placement of *Spirodactylon aureum* as well as the type species of *Coemansia* and *Kickxella*. The red star highlights the branch subtending the insect-associated groups (Asellariales, Barbatosporales, Harpellales, Orphellales). Images depict examples of host organisms of selected clades.

Figure 2. Genome assembly summary for the low coverage genome data, including number of contigs in the assembly, total length of the assembly (in megabases, Mb), the number of fragmented and complete genes from BUSCO analyses, average coverage, number of predicted genes from the annotation, and number of marker genes recovered by PHYling out of a total of 758.

Figure 3. Correlation matrix showing pairwise Pearson correlation values for the relationships between average coverage, contig count, BUSCO scores (percent complete, fragmented, and duplicated), number of predicted genes, and number of marker genes recovered by PHYling out of a total of 758. The correlation coefficient is shown and the asterisks indicate significant values (*** = p < 0, ** = p < 0.001, * p <

0.01, . = p < 0.05). The red line indicates the regression line with 95% confidence intervals in blue. Graphs along the diagonal show the distribution of each variable. *Coemansia* sp. IMI 209128 was an outlier in the average coverage (827.5x coverage) so it was removed prior to analyses.

Figure 4. ASTRAL coalescent-based species tree compiled from 461 RAxML gene trees. Major clades are indicated by boxes and isolate names in bold are type cultures. Isolates in black text are high quality reference genomes downloaded from the Joint Genome Institute while those in blue text are low coverage genomes. Red arrows indicate the type species for the genera *Coemansia* and *Kickxella*. All unlabeled branches are fully supported. Because ASTRAL analyzes quartets, each node has three possible topologies. Pie charts at selected nodes represent the proportion of gene trees with the main topology (q1), alternate topology 1 (q2), and alternate topology 2 (q3).

Figure 5. OrthoFinder phylogeny based on 101 orthogroups aligned with MUSCLE and reconstructed with FastTree Approximately ML. All unlabeled branches are fully supported. Isolates in black text are high quality reference genomes downloaded from the Joint Genome Institute while those in blue text are low coverage genomes. Red arrows indicate the type species for the genera *Coemansia* and *Kickxella*.

Figure 6. A) Box plots of predicted secondary metabolites and selected proteases for the major clades of Kickxellomycotina. B) Bar graphs showing the numbers of genes

annotated to various CAZyme families for each of the major clades. Note that the y axis scale differs for each plot in A and B and that siderophores are mostly encoded by NRPS domains. CBM = carbohydrate binding module.

Figure 7. Linear regressions for the number of CAZymes (as predicted by InterProScan against the dbCAN database) versus the number of proteases (as predicted by InterProScan against the MEROPS database) for low coverage genome data by clade. The R^2 value is given for each association and bolded values are statistically significant (p value codes: 0 '***', 0.001 '**', 0.01 '*').

REFERENCES

Ahrendt, S.R., Quandt, C.A., Ciobanu, D., Clum, A., Salamov, A., Andreopoulos, B., Cheng, J.F., Woyke, T., Pelin, A., Henrissat, B., Reynolds, N.K., Benny, G.L., Smith, M.E., James, T.Y., Grigoriev, I.V. Leveraging single-cell genomics to expand the fungal tree of life. Nature Microbiology, 3(12): 1417-1428. doi: 10.1038/s41564-018-0261-0

Albertin, W., Marullo, P. 2012. Polyploidy in fungi: evolution after whole-genome duplication. Proc. R. Soc. B. 279: 2497-2509. http://doi.org/10.1098/rspb.2012.0434

Allen, J.M., Huang, D.I., Cronk, Q.C., Johnson, K.P. 2015. aTRAM - automated target restricted assembly method: a fast method for assembling loci across divergent taxa from next-generation sequencing data. BMC Bioinformatics, 16: 98. https://doi.org/10.1186/s12859-015-0515-2

Alonge, M., Soyk, S., Ramakrishnan, S. Wang, X., Goodwin, S., Sedlazeck, F.J., Lippman, Z.B., Schatz, M.C. 2019. RaGOO: fast and accurate reference-guided scaffolding of draft genomes. Genome Biol., 20: 224. https://doi.org/10.1186/s13059-019-1829-6

Amses, K.R., D. Simmons, R., Longcore, J.E., Mondo, S.J., Seto, K., Jerônimo, G.H., Bondsa, A.E., Quandt, C.A., Davis, W.J., Chang, Y., Federici, B.A., Kuo, A., LaButti, K., Pangilinan, J., Andreopoulos, W., Tritt, A., Riley, R., Hundley, H., Johnson, J., Lipzen, A., Barry, K., Lang, B.F., Cuomo, C.A., Buchler, N.E., Grigoriev, I.V., Spatafora, J.W., Stajich, J.E., James, T.Y. 2022. Diploid-dominant life cycles characterize the early evolution of Fungi. PNAS, 119(36): e2116841119. https://doi.org/10.1073/pnas.2116841119

Baldrian, P. 2005. Fungal laccases – occurrence and properties. FEMS Microbiology Reviews, 30: 215–242. doi: 10.1111/j.1574-4976.2005.00010.x

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M. Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology, 19(5): 455–477. doi: 10.1089/cmb.2012.0021

Battaglia, E., Benoit, I., van den Brink, J., Wiebenga, A., Coutinho, P.M., Henrissat, B., de Vries, R.P. 2011. Carbohydrate-active enzymes from the zygomycete fungus *Rhizopus oryzae*: a highly specialized approach to carbohydrate degradation depicted at genome level. BMC Genomics, 12: 38. http://www.biomedcentral.com/1471-2164/12/38

Beidler, K.V., Phillips, R.P., Andrews, E., Maillard, F., Mushinski, R.M., Kennedy, P.G. 2020. Substrate quality drives fungal necromass decay and decomposer community structure under contrasting vegetation types. Journal of Ecology, 108: 1845–1859. doi: 10.1111/1365-2745.13385

Benítez, T., Rincón, A.M., Limón, M.C., Codón, A.C. 2004. Biocontrol mechanisms of *Trichoderma* strains. International Microbiology, 7: 249-260.

Benjamin, R.K. 1958. Sexuality in the Kickxellaceae. Aliso, 4: 149-169.

Benjamin, R.K. 1959. The merosporangiferous Mucorales. Aliso, 4: 321-433.

Benjamin, R.K. 1961. Addenda to “The merosporangiferous Mucorales”. Aliso, 5(1): 11-19.

Benjamin, R.K. 1963. Addenda to “The merosporangiferous Mucorales” II. Aliso 5:273-288.

Benny, G.L. 2008. Methods Used by Dr. R. K. Benjamin, and Other Mycologists, to Isolate Zygomycetes. Aliso, 26(1): 37-61.

Benny, G.L., Smith, M.E., Kirk, P.M., Tretter, E.D., White, M.M. 2016. Challenges and future perspectives in the systematics of Kickxellomycotina, Mortierellomycotina, Mucoromycotina, and Zoopagomycotina. In: Li D-W, ed. Biology of microfungi. Cham, Switzerland: Springer International Publishing. pp 65–126.

Blin, K., Shaw, S., Steinke, K., Villebro, R., Ziemert, N., Lee, S.Y., Medema, M.H., Weber, T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. Nucleic Acids Research, 47: W81-W87.

Bolger, A. M., Lohse, M., Usadel, B. 2014. Trimmomatic: A flexible trimmer for Illumina Sequence Data, Bioinformatics, btu170.

Bonito, G., Smith, M.E., Nowak, M., Healy, R.A., Guevara, G., Cázares, E., Kinoshita, A., Nouhra, E.R., Domínguez, L.S., Tedersoo, L., Murat, C., Wang, Y., Moreno, B.A., Pfister, D.H., Nara, K., Zambonelli, A., Trappe, J.M., Vilgalys, R. 2013. Historical Biogeography and Diversification of Truffles in the Tuberaceae and Their Newly Identified Southern Hemisphere Sister Lineage. PLoS ONE, 8(1): e52765. doi: 10.1371/journal.pone.005276

Bonugli-Santos, R.C., Durrant, L.R., da Silva, M., Durães Sette, L. 2009. Production of laccase, manganese peroxidase and lignin peroxidase by Brazilian marine-derived fungi. Enzyme and Microbial Technology, 46: 32–37. doi: 10.1016/j.enzmictec.2009.07.014

Brouta, F., Descamps, F., Monod, M., Vermout, S., Losson, B., Mignon, B. 2002. Secreted Metalloprotease Gene Family of *Microsporum canis*. Infection and Immunity, 70(10): 5676–5683. doi: 10.1128/IAI.70.10.5676–5683.2002

Brown, C.T., Irber, L. 2016. sourmash: a library for MinHash sketching of DNA. Journal of Open Source Software, 1(5): 27. doi:10.21105/joss.00027

Buchfink, B., Reuter, K., Drost, H.G. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nature Methods, 18: 366–368. doi:10.1038/s41592-021-01101-x

Bushley, K.E., Turgeon, B.G. 2010. Phylogenomics reveals subfamilies of fungal nonribosomal peptide synthetases and their evolutionary relationships. BMC Evolutionary Biology, 10: 26. http://www.biomedcentral.com/1471-2148/10/26

Bushnell, B. 2014. BBMap: A Fast, Accurate, Splice-Aware Aligner. Lawrence Berkeley National Laboratory. LBNL Report #: LBNL-7065E. Retrieved from https://escholarship.org/uc/item/1h3515gn

Challis, R., Richards, E., Rajan, J., Cochrane, G., Blaxter, M. 2020. BlobToolKit – Interactive Quality Assessment of Genome Assemblies, G3 Genes|Genomes|Genetics, 10(4): 1361–1374. https://doi.org/10.1534/g3.119.400908

Chang, Y. 1967. Linderina macrospora sp. nov. from Hong Kong. Transactions of the British Mycological Society, 50: 311-314.

Chang, Y., Desiró, A., Na, H., Sandor, L., Lipzen, A., Clum, A., Barry, K., Grigoriev, I.V., Martin, F.M., Stajich, J.E., Smith, M.E., Bonito, G, Spatafora, J.W. 2018. Phylogenomics of Endogonaceae and evolution of mycorrhizas within Mucoromycota. New Phytologist, 222: 511–525. doi: 10.1111/nph.15613

Chang, Y., Wang, S., Sekimoto, S., Aerts, A.L., Choi, C., Clum, A., LaButti, K.M., Lindquist, E.A., Yee Ngan, C., Ohm, R.A., Salamov, A.A., Grigoriev, I.V., Spatafora, J.W., Berbee, M.L. 2015. Phylogenomic Analyses Indicate that Early Fungi Evolved Digesting Cell Walls of Algal Ancestors of Land Plants. Genome Biology and Evolution, 7(6): 1590-601. doi: 10.1093/gbe/evv090

Chuang, S.-C., Ho, H.-M., Reynolds, N., Smith, M.E., Benny, G.L., Chien, C.-Y., Tsai, J.-L. 2017. Preliminary phylogeny of *Coemansia* (Kickxellales), with descriptions of four new species from Taiwan. Mycologia, 109(5): 815-831. doi: 10.1080/00275514.2017.1401892

Coemans, E. 1862. Spicilége mycologique. No. III. Notice sur un champignon nouveau: *Kickxella alabastrina*. Cms. Bulletin de la Société Royale de Botanique de Belgique, 1: 155-159.

Corrochano, L.M., Kuo, A., Marcet-Houben, M., Polaino, S., Salamov, A., Villalobos-Escobedo, J.M., Grimwood, J., Álvarez, M.I., Avalos, J., Bauer, D., Benito, E.P., Benoit, I., Burger, G., Camino, L.P., Cánovas, D., Cerdá-Olmedo, E., Cheng, J-F., Domínguez, A., Eliáš, M., Eslava, A.P., Glaser, F., Gutiérrez, G., Heitman, J., Henrissat, B., Iturriaga, E.A., Lang, B.F., Lavín, J.L., Lee, S.C., Li, W., Lindquist, E., López-García, S., Luque, E.M., Marcos, A.T., Martin, J., McCluskey, K., Medina, H.R., Miralles-Durán, A., Miyazaki, A., Muñoz-Torres, E., Oguiza, J.A., Ohm, R.A., Olmedo, M., Orejas, M., Ortiz-Castellanos, L., Pisabarro, A.G., Rodríguez-Romero, J., Ruiz-Herrera, J., Ruiz-Vázquez, R., Sanz, C., Schackwitz, W., Shahriari, M., Shelest, E., Silva-Franco, F., Soanes, D., Syed, K., Tagua, V.G., Talbot, N.J., Thon, M.R., Tice, H., de Vries, R.P., Wiebenga, A., Yadav, J.S., Braun, E.L., Baker, S.E., Garre, V., Schmutz, J., Horwitz, B.A., Torres-Martínez, S., Idnurm, A., Herrera-Estrella, A., Gabaldón, T., Grigoriev, I.V. Expansion of Signal Transduction Pathways in Fungi by Extensive Genome Duplication. Current Biology, 26(12): 1577-1584. https://doi.org/10.1016/j.cub.2016.04.038

Cox R.J., Simpson, T.J. 2009. Fungal type I polyketide synthases. In Abelson, J.A., Simon, M.I., Colowick, S.P., Kaplan N.O. (Eds.), Methods in Enzymology, 459:49-78. doi: 10.1016/S0076-6879(09)04603-5

Davies, G., Attia, M. 2021. "Glycoside Hydrolase Family 5" in CAZypedia, available at URL http://www.cazypedia.org/, accessed 11 February 2022.

Davis, W.J., Amses, K.R., Benny, G.L., Carter-House, D., Chang, Y., Grigoriev, I.V., Smith, M.E., Spatafora, J.W., Stajich, J.E., James, T.Y. 2019. Genome-scale phylogenetics reveals a monophyletic Zoopagales (Zoopagomycota, Fungi). Molecular Phylogenetics and Evolution, 133: 152-163. doi: 10.1016/j.ympev.2019.01.006

Deveau, A., Bonito, G., Uehling, J., Paoletti, M., Becker, M., Bindschedler, S., Hacquard, S., Hervé, V., Labbé, J., Lastovetsky, O.A., Mieszkin, S., Millet, L.J., Vajna, B., Junier, P., Bonfante, P., Krom, B.P., Olsson, S., van Elsas, J.D., Wick, L.Y. 2018. Bacterial-fungal interactions: ecology, mechanisms and challenges. FEMS Microbiology Reviews, 42(3): 335-352. doi: 10.1093/femsre/fuy008

Doweld, A.B. 2014. Nomenclatural novelties [Ramicandelaberales]. Index Fungorum no. 69: 1

Eddy, S.R. 2011. Accelerated profile HMM searches. PLOS Comp. Biol.,7: e1002195.

Eddy, S.R. and the HMMER development team. 2020. HMMER User's Guide version 3.3.2. http://hmmer.org

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32(5): 1792-1797.

Eidam, E. 1887. *Coemansia spiralis*. Jahres-Bericht der Schlesischen Gesellschaft für Vaterländische Cultur, 65: 262-263.

Emms, D.M., Kelly, S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biology, 16: 157. https://doi.org/10.1186/s13059-015-0721-2

Emms, D.M., Kelly, S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology, 20: 238. https://doi.org/10.1186/s13059-019-1832-y

Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R., Heger, A., Hetherington, K., Holm, L., Mistry, J., Sonnhammer, E. L., Tate, J., Punta, M. 2014. Pfam: the protein families database. Nucleic Acids Research, 42(Database issue): D222–D230. https://doi.org/10.1093/nar/gkt1223

Floudas,D., Binder, M., Riley, R., Barry, K., Blanchette, R.A., Henrissat, B., Martínez, A.T., Otillar, R., Spatafora, J.W., Yadav, J.S., Aerts, A., Benoit, I., Boyd, A., Carlson, A., Copeland, A., Coutinho, P.M., de Vries, R.P., Ferreira, P., Findley, K., Foster, B., Gaskell, J., Glotzer, D., Górecki, P., Heitman, J., Hesse, C., Hori, C., Igarashi, K.,

Jurgens, J.A., Kallen, N., Kersten, P., Kohler, A., Kües, U., Kumar, T.K.A., Kuo, A., LaButti, K., Larrondo, L.F., Lindquist, E., Ling, A., Lombard, V., Lucas, S., Lundell, T., Martin, R., McLaughlin, D.J., Morgenstern, I., Morin, E., Murat, C., Nagy, L.G., Nolan, M., Ohm, R.A., Patyshakuliyeva, A., Rokas, A., Ruiz-Dueñas, F.J., Sabat, G., Salamov, A., Samejima, M., Schmutz, J., Slot, J.C., St. John, F., Stenlid, J., Sun, H., Sun, S., Syed, K., Tsang, A., Wiebenga, A., Young, D., Pisabarro, A., Eastwood, D.C., Martin, F., Cullen, D., Grigoriev, I.V., Hibbett, D.S. 2012. The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. Science, 336: 1715-1719. doi: 10.1126/science.1221748

Freitas, A.C., Ferreira, F., Costa, A.M., Pereira, R., Antunes, A.C., Gonçalves, F., Rocha-Santos, A.P., Diniz, M.S., Castro, L., Peres, I., Duarte, A.C. Biological treatment of the effluent from a bleached kraft pulp mill using basidiomycete and zygomycete fungi. Science of the Total Environment, 407: 3282–3289. doi: 10.1016/j.scitotenv.2009.01.054

Galindo, L.J., López-García, P., Torruella, G., Karpov, S., Moreira, D. 2021. Phylogenomics of a new fungal phylum reveals multiple waves of reductive evolution across Holomycota. bioRxiv 2020.11.19.389700; doi: 10.1101/2020.11.19.389700

Gawad, C., Koh, W., Quake, S.R. 2016. Single-cell genome sequencing: current state of the science. Nature Reviews Genetics, 17: 175-188. doi:10.1038/nrg.2015.16

Grabherr, M., Haas, B., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol 29, 644–652. https://doi.org/10.1038/nbt.1883

Grigoriev, I.V., Nikitin, R., Haridas, S., Kuo, A., Ohm, R., Otillar, R., Riley, R., Salamov, A., Zhao, X., Korzeniewski, F., Smirnova, T., Nordberg, H., Dubchak, I., Shabalov, I. 2014. MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Research, 42(D1): D699–D704. https://doi.org/10.1093/nar/gkt1183

Grube, M., Wedin, M. 2016. Lichenized Fungi and the Evolution of Symbiotic Organization. Microbiology Spectrum, 4(6). doi: 10.1128/microbiolspec.FUNK-0011-2016

Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., Wortman, J.R. 2008. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. Genome Biol., 9: R7. https://doi.org/10.1186/gb-2008-9-1-r7

Hage, H., Rosso, M.-N. 2021. Evolution of Fungal Carbohydrate-Active Enzyme Portfolios and Adaptation to Plant Cell-Wall Polymers. Journal of Fungi, 7: 185. https://doi.org/10.3390/jof7030185

Hartl, L., Zach, S., Seidl-Seiboth, V. 2012. Fungal chitinases: diversity, mechanistic properties and biotechnological potential. Applied Microbiology and Biotechnology, 93: 533–543. doi: 10.1007/s00253-011-3723-3

Hawksworth, D.L., Lücking, R. 2017. Fungal diversity revisited: 2.2 to 3.8 million species. Microbiol Spectrum, 5(4):FUNK-0052-2016. doi: 10.1128/microbiolspec.FUNK-0052-2016.

Hibbett, D.S., Binder, M., Bischoff, J.F., Blackwell, M., Cannon, P.F., Eriksson, O.E., Huhndorf, S., James, T., Kirk, P.M., Lücking, R., Lumbsch, H., Lutzoni, F., Matheny, P.B., McLaughlin, D.J., Powell, M.J., Redhead, S., Schoch, C.L., Spatafora, J.W., Stalpers, J.A., Vilgalys, R., Aime, M.C., Aptroot, A., Bauer, R., Begerow, D., Benny, G.L., Castlebury, L.A., Crous, P.W., Dai, Y.C., Gams, W., Geiser, D.M., Griffith, G.W., Gueidan, C., Hawksworth, D.L., Hestmark, G., Hosaka, K., Humber, R.A., Hyde, K.D., Ironside, J.E., Kõljalg, U., Kurtzman, C.P., Larsson, K.H., Lichtwardt, R.W., Longcore, J., Miądikowska, J., Miller, A., Moncalvo, J.M., Mozley-Standridge, S., Oberwinkler, F., Parmasto, E., Reeb, V., Rogers, J.D., Roux, C., Ryvarden, L., Sampaio, J.P., Schüssler, A., Sugiyama, J., Thorn, R.G., Tibell, L., Untereiner, W.A., Walker, C., Wang, Z., Weir, A., Weiss, M., White, M.M., Winka, K., Yao, Y.J., Zhang, N. 2007. A higher-level phylogenetic classification of the Fungi. Mycological Research, 111: 509–547. doi: 10.1016/j.mycres.2007.03.004

Huang, X., Madan, A. 1999. CAP3: A DNA sequence assembly program. Genome Research, 9: 868-877.

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., von Mering, C., Bork, P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Research, 47(D1): D309–D314. https://doi.org/10.1093/nar/gky1085

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

Inglis, D.O., Binkley, J., Skrzypek, M.S., Arnaud, M.B., Cerqueira, G.C., Shah, P., Wymore, F., Wortman, J.R., Sherlock, G. 2013. Comprehensive annotation of secondary metabolite biosynthetic genes and gene clusters of *Aspergillus nidulans, A. fumigatus, A. niger* and *A. oryzae*. BMC Microbiol, 13(91). https://doi.org/10.1186/1471-2180-13-91

Iqbal, M., Dubey, M., Gudmundsson, M., Viketoft, M., Jensen, D.F., Karlsson, M. 2018. Comparative evolutionary histories of fungal proteases reveal gene gains in the mycoparasitic and nematode-parasitic fungus *Clonostachys rosea*. BMC Evolutionary Biology, 18: 171. doi: 10.1186/s12862-018-1291-1

Jackson, H.S., Dearden, E.R. 1948. *Martensella corticii* Thaxter and its distribution. Mycologia 40:168-176.

Janusz, G., Pawlik, A., Sulej, J., Świderska-Burek, U., Jarosz-Wilkołazka, A., Paszczyński, A. 2017. Lignin degradation: microorganisms, enzymes involved, genomes analysis and evolution. FEMS Microbiology Reviews, 41: 941–962. doi: 10.1093/femsre/fux049

Kempken, F., Rohlfs, M. 2009. Fungal secondary metabolite biosynthesis – a chemical defense strategy against antagonistic animals? Fungal Ecology, 3: 107-114. doi:10.1016/j.funeco.2009.08.001

Kurihara, Y., Degawa, Y., Tokumasu, S. 2004. Two novel kickxellalean fungi, *Mycoëmilia scoparia* gen. sp. nov. and *Ramicandelaber brevisporus* sp. nov. Mycological Research, 108(10): 1143–1152. doi: 10.1017/S0953756204000930

Kurihara, Y., Sukarno, N., Ilyas, M., Yuniarti, E., Mangunwardoyo, W., Park, J.-Y., Saraswati, R., Widyastuti, Y., Ando, K. 2008. Indonesian Kickxellales: two species of *Coemansia* and *Linderina*. Mycoscience, 49: 250–257. doi: 10.1007/s10267-008-0417-5

Kwaśna, H., Richardson, M.J., Bateman, G.L. 2002. Morphological variation in the *Coemansia spiralis* complex. Mycological Research, 106(2): 252-256. doi: 10.1017/S0953756201005421

Laczny, C.C., Sternal, T., Plugaru, V., Gawron, P., Atashpendar, A., Margossian, H.H., Coronado, S., van der Maaten, L., Vlassis, N., Wilmes, P. 2015. VizBin - an application for reference-independent visualization and human-augmented binning of metagenomic data. Microbiome, 3(1): 1. doi:10.1186/s40168-014-0066-1

Laetsch, D.R., Blaxter, M.L. 2017. KinFin: Software for Taxon-Aware Analysis of Clustered Protein Sequences. G3: Genes, Genomes, Genetics, 7(10): 3349–3357. doi:10.1534/g3.117.300233

Lang, L., Barrett, K., Pilgaard, B., Gleason, F., Tsang, A. 2019. Enzymes of early-diverging, zoosporic fungi. Applied Microbiology and Biotechnology, 103: 6885–6902. https://doi.org/10.1007/s00253-019-09983-w

Latgé, J.-P. 2007. The cell wall: a carbohydrate armour for the fungal cell. Molecular Microbiology, 66(2): 279–290. doi:10.1111/j.1365-2958.2007.05872.x

Li, H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34:18, 3094–3100. https://doi.org/10.1093/bioinformatics

Li, J., Gu, F., Wu, R., Yang, J., Zhang, K.-Q. 2017. Phylogenomic evolutionary surveys of subtilase superfamily genes in fungi. Scientific Reports, 7: 45456. doi: 10.1038/srep45456

Lichtwardt, R.W., Cafaro, M.J., White, M.M. 2001. The Trichomycetes, fungal associates of arthropods. Revised edition. [updated 2018 Jan; cited 2021 July 21]. Available from: https://keyserver.lucidcentral.org/key-server/data/0b08020c-0f0c-4908-8807-030c020a0002/media/Html/monograph/text/mono.htm

Lilly, W.W., Stajich, J.E., Pukkila, P.J., Wilke, S.K., Inoguchi, N., Gathman, A.C. 2007. An expanded family of fungalysin extracellular metallopeptidases of Coprinopsis cinerea. Mycological Research, 112: 389–398. doi: 10.1016/j.mycres.2007.11.013

Ma, L-J., Ibrahim, A.S., Skory, C., Grabherr, M.G., Burger, G., Butler, M., Elias, M., Idnurm, A., Lang, B.F., Sone, T., Abe, A., Calvo, S.E., Corrochano, L.M., Engels, R., Fu, J., Hansberg, W., Kim, J-M., Kodira, C.D., Koehrsen, M.J., Liu, B., Miranda-Saavedra, D., O'Leary, S., Ortiz-Castellanos, L., Poulter, R., Rodriguez-Romero, J., Ruiz-Herrera, J., Shen, Y-Q., Zeng, Q., Galagan, J., Birren, B.W., Cuomo , C.A., Wickes, B.L. 2009. Genomic Analysis of the Basal Lineage Fungus Rhizopus oryzae Reveals a Whole-Genome Duplication. PLoS Genet 5(7): e1000549. doi: 10.1371/journal.pgen.1000549

Macheleidt, J., Mattern, D.J., Fischer, J., Netzker, T., Weber, J., Schroeckh, V., Valiante, V., Brakhage, A.A. 2016. Regulation and Role of Fungal Secondary Metabolites. Annual Review of Genetics, 50: 371–392. doi: 10.1146/annurev-genet-120215-035203

Markaryan, A., Morozova, I., Yu, H., Kolaitukudy, P.E. 1994. Purification and Characterization of an Elastinolytic Metalloprotease from *Aspergillus fumigatus* and Immunoelectron Microscopic Evidence of Secretion of This Enzyme by the Fungus Invading the Murine Lung. Infection and Immunity, 62(6): 2149-2157.

Meyer, J. 1957. Martensiomyces pterosporus nov. gen. nov. sp. nouvelle Kickxellacée isolee du sol. Bulletin de la Société mycologique de France, 73: 189-201.

Miller, I.J., Rees, E.R., Ross, J., Miller, I., Baxa, J., Lopera, J., Kerby, R.L., Rey, F.E., Kwan, J.C. 2019. Autometa: Automated extraction of microbial genomes from individual shotgun metagenomes. Nucleic Acids Research, 47(10): e57. https://doi.org/10.1093/nar/gkz148

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., Lanfear, R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol., 37:1530-1534. https://doi.org/10.1093/molbev/msaa015

Modlin, S.J., Robinhold, C., Morrissey, C., Mitchell, S.N., Ramirez-Busby, S.M., Shmaya, T., Valafar, F. 2021. Exact mapping of Illumina blind spots in the Mycobacterium tuberculosis genome reveals platform-wide and workflow-specific biases. Microbial Genomics, 7:000465. doi: 10.1099/mgen.0.000465

Mondo, S.J., Dannebaum, R.O., Kuo, R.C., Louie, K.B., Bewick, A.J., LaButti, K., Haridas, S., Kuo, A., Salamov, A., Ahrendt, S.R., Lau, R., Bowen, B.P., Lipzen, A., Sullivan, W., Andreopoulos, B.B., Clum, A., Lindquist, E., Daum, C., Northen, T.R., Kunde-Ramamoorthy, G., Schmitz, R.J., Gryganskyi, A., Culley, D., Magnuson, J., James, T.Y., O'Malley, M.A., Stajich, J.E., Spatafora, J.W., Visel, A., Grigoriev, I.V. 2017. Widespread adenine N6-methylation of active genes in fungi. Nature Genetics, 49(6): 964-968. doi: 10.1038/ng.3859

Muszewska, A., Taylor, J.W., Szczesny, P., Grynberg, M. 2011. Independent Subtilases Expansions in Fungi Associated with Animals. Molecular Biology and Evolution, 28(12): 3395–3404. doi:10.1093/molbev/msr176

Ogawa, Y., Hayashi, S., Degawa, Y., Yaguchi, Y. 2001. *Ramicandelaber*, a new genus of the Kickxellales, Zygomycetes. Mycoscience, 42: 193-199.

Ökmen, B., Kemmerich, B., Hilbig, D., Wemhöner, R., Aschenbroich, J., Perrar, A., Huesgen, P.F., Schipper, K., Doehlemann, G. 2018. Dual function of a secreted fungalysin metalloprotease in *Ustilago maydis*. New Phytologist, 220: 249–261. doi: 10.1111/nph.15265

Oliveros, J.C. 2007-2015. Venny: An interactive tool for comparing lists with Venn's diagrams. https://bioinfogp.cnb.csic.es/tools/venny/index.html

Palmer, J., Stajich, J.E. 2020. nextgenusfs/funannotate: funannotate v1.8.1 (Version 1.8.1). Zenodo. https://doi.org/10.5281/zenodo.4054262

Payne, C.M., Knott, B.C., Mayes, H.B., Hansson, H., Himmel, M.E., Sandgren, M., Ståhlberg, J., Beckham, G.T. 2015. Fungal Cellulases. Chemical Reviews, 115: 1308−1448. doi: 10.1021/cr500351c

Petkovits, T., Nagy, L.G., Hoffmann, K., Wagner, L., Nyilasi, I., Griebel, T., Schnabelrauch, D., Vogel, H., Voigt, K., Vágvölgyi, C., Papp, T. 2011. Data Partitions, Bayesian Analysis and Phylogeny of the Zygomycetous Fungal Family Mortierellaceae, Inferred from Nuclear Ribosomal DNA Sequences. PLOS ONE 6(11): e27507. https://doi.org/10.1371/journal.pone.0027507

Phillips, M.J., Penny, D. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Molecular Phylogenetics and Evolution, 28(2): 171-185.

Pinard, R., de Winter, A., Sarkis, G.J., Gerstein, M.B., Tartaro, K.R., Plant, R.N., Egholm, M., Rothberg, J.M., Leamon, J.H. 2006. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics, 7: 216. doi:10.1186/1471-2164-7-216

Pombejra, S.N., Jamklang, M., Uhrig, J.P., Vu, K., Gelli, A. 2018. The structure-function analysis of the Mpr1 metalloprotease determinants of activity during migration of fungal

cells across the blood-brain barrier. PLoS ONE, 13(8): e0203020. https://doi.org/10.1371/journal.pone.0203020

Price, M.N., Dehal, P.S., Arkin, A.P. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE, 5(3): e9490. https://doi.org/10.1371/journal.pone.0009490

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., Korobeynikov, A. 2020. Using SPAdes de novo assembler. Current Protocols in Bioinformatics, 70: e102. doi:10.1002/cpbi.102

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raper, K.B., Fennell, D.I. 1952. Two noteworthy fungi from Liberian soil. American Journal of Botany, 39: 79-86.

Rawlings, N.D., Barrett, A.J., Thomas, P.D., Huang, X., Bateman, A., Finn, R.D. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. Nucleic Acids Res., 46: D624-D632.

Rhodes J., Beale M.A., Fisher M.C. 2014. Illuminating Choices for Library Prep: A Comparison of Library Preparation Methods for Whole Genome Sequencing of *Cryptococcus neoformans* Using Illumina HiSeq. PLoS ONE, 9(11): e113501. doi: 10.1371/journal.pone.0113501

Ri, T., Suyama, M., Takashima, Y., Seto, K., Yousuke, D. 2022. A new genus *Unguispora* in Kickxellales shows an intermediate lifestyle between saprobic and gut-inhabiting fungi. Mycologia, 114(6): 934-946. doi: 10.1080/00275514.2022.2111052

Robinson, S.L., Christenson, J.K., Wackett, L.P. 2019. Biosynthesis and chemical diversity of β-lactone natural products. Natural Products Reports, 36: 458. doi: 10.1039/c8np00052b

Rokas, A., Wisecaver, J.H., Lind, A.L. 2018. The birth, evolution and death of metabolic gene clusters in fungi. Nature Reviews Microbiology, 16: 731-744. doi: 10.1038/ s41579-018-0075-3

Rosenblum, E.B., Stajich, J.E., Maddox, N., Eisen, M.B. 2008. Global gene expression profiles for life stages of the deadly amphibian pathogen *Batrachochytrium dendrobatidis*. PNAS, 105(44): 17034–17039. doi: 10.1073 pnas.0804173105

Sabina J., Leamon J.H. 2015. Bias in Whole Genome Amplification: Causes and Considerations. In: Kroneis T ed. Whole Genome Amplification: Methods in Molecular

Biology, vol 1347. Humana Press, New York, NY. pp 15-41. https://doi.org/10.1007/978-1-4939-2990-0_2

Sanz-Martín, J.M., Pacheco-Arjona, J.R., Bello-Rico, V., Vargas, W.A., Monod, M., Díaz-Mínguez, J.M., Thon, M.R., Sukno, S.A. 2016. Molecular Plant Pathology, 17(7): 1048–1062. doi: 10.1111/mpp.12347

Sato, M.P., Ogura, Y., Nakamura, K., Nishida, R., Gotoh, Y., Hayashi, M., Hisatsune, J., Sugai, M., Takehiko, I., Hayashi, T. 2019. Comparison of the sequencing bias of currently available library preparation kits for Illumina sequencing of bacterial genomes and metagenomes. DNA Research, 26(5): 391–398. doi: 10.1093/dnares/dsz017

Sayyari, E., Mirarab, S. 2016. Fast Coalescent-Based Computation of Local Branch Support from Quartet Frequencies. Molecular Biology and Evolution, 33(7): 1654–68. doi:10.1093/molbev/msw079

Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A., Crowley, J. 2021. Package 'GGally' version 2.1.2. The R Project for Statistical Computing. doi: 10.5281/zenodo.5009047

Seidl, V. 2008. Chitinases of filamentous fungi: a large group of diverse proteins with multiple physiological functions. Fungal Biology Reviews, 22: 36-42. doi:10.1016/j.fbr.2008.03.002

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics, 31(19): 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Slater, G.S.C., Birney, E. 2005. Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics, 6: 31. https://doi.org/10.1186/1471-2105-6-31

Smit, A.F.A., Hubley, R., Green, P. 2013-2015. RepeatMasker Open-4.0. http://www.repeatmasker.org

Sørensen, A., Lübeck, M., Lübeck, P.S., Ahring, B.K. 2013. Fungal Beta-Glucosidases: A Bottleneck in Industrial Use of Lignocellulosic Materials, Biomolecules, 3: 612-631. doi:10.3390/biom3030612

Stajich, J.E., Gilchrist, C., Palmer J.M. 2021. Stajichlab/AAFTF: v.0.2.5 release with (Version v.0.2.5). Zenodo. http://doi.org/10.5281/zenodo.4643064

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

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., Morgenstern, B. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res., 34(Web Server issue): W435-9. doi: 10.1093/nar/gkl200

Steenwyk, J.L., Buida III, T.J., Labella, A.L., Li, Y., Shen, X.-X., Rokas, A. 2021. PhyKIT: a broadly applicable UNIX shell toolkit for processing and analyzing phylogenomic data. Bioinformatics, btab096. doi: 10.1093/bioinformatics/btab096

Steenwyk, J.L., Buida III, T.J., Li, Y., Shen, X.-X., Rokas, A. 2020. ClipKIT: A multiple sequence alignment trimming software for accurate phylogenomic inference. PLoS Biol., 18(12): e3001007. doi: 10.1371/journal.pbio.3001007

Strassert, J.F.H., Monaghan, M.T. 2022. Phylogenomic insights into the early diversification of fungi. Current Biology, 32: 1–8. doi: 10.1016/j.cub.2022.06.057

Ter-Hovhannisyan, V., Lomsadze, A., Chernoff, Y.O., Borodovsky, M. 2008. Gene prediction in novel fungal genomes using an ab initio algorithm with unsupervised training. Genome Res., 18(12): 1979–1990.

The UniProt Consortium. 2021. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research, 49(D1): D480–D489. https://doi.org/10.1093/nar/gkaa1100

Valle, L.G., Cafaro, M.J. 2008. First report of Zygospores in Asellariales and new species from the Caribbean. Mycologia, 100(1): 122-131.

van Tieghem, P., and G. Le Monnier. 1873. Recherches sur les Mucorinées. Annales des Sciences Naturelles, Botanique, Séries V, 17: 261-399.

Vandepol, N., Liber, J., Desirò, A., Na, H., Kennedy, M., Berry, K., Grigoriev, I.V., Miller, A.N., O'Donnell, K., Stajich, J.E., Bonito, G. 2020. Resolving the Mortierellaceae phylogeny through synthesis of multi-gene phylogenetics and phylogenomics. Fungal Diversity, 104: 267-289. doi: 10.1007/s13225-020-00455-5

Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., Earl, A.M. 2014. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. PLOS ONE, 9(11): e112963. https://doi.org/10.1371/journal.pone.0112963

Wang, Y., Tretter, E.D., Lichtwardt, R.W., White, M.M. 2013. Overview of 75 years of *Smittium* research, establishing a new genus for *Smittium culisetae*, and prospects for future revisions of the 'Smittium' clade, Mycologia, 105:1, 90-111. doi: 10.3852/11-311

Wang, Y., White, M.M., Kvist, S., Moncalvo, J.-M. 2016. Genome-Wide Survey of Gut Fungi (Harpellales) Reveals the First Horizontally Transferred Ubiquitin Gene from a

Mosquito Host. Molecular Biology and Evolution, 33(10): 2544–2554. doi: 10.1093/molbev/msw126

Wang, Y., White, M.M., Moncalvo, J.-M. 2016b. Draft Genome Sequence of *Capniomyces stellatus*, the Obligate Gut Fungal Symbiont of Stonefly. Genome Announcement, 4(4): e00804-16. doi:10.1128/genomeA.00761-16

Wang, Y., Stata, M., Wang, W., Stajich, J.E., White, M.M., Moncalvo, J-M. 2018. Comparative genomics reveals the core gene toolbox for the fungus-insect symbiosis. mBio, 9: e00636-18. https://doi.org/10.1128/mBio.00636-18.

White, M.M., Guàrdia Valle, L., Lichtwardt, R.W., Siri, A., Strongman, D.B., William, R.T., Gause, W.J., Tretter, E.D. 2018. New species and emendations of *Orphella*: taxonomic and phylogenetic reassessment of the genus to establish the Orphellales, for stonefly gut fungi with a twist. Mycologia, 110(1): 147-178. doi: 10.1080/00275514.2018.1448198

White, M.M., James, T.Y., O'Donnell, K., Cafaro, M.J., Tanabe, Y., Sugiyama, J. 2006b. Phylogeny of the Zygomycota Based on Nuclear Ribosomal Sequence Data. Mycologia, 98(6): 872-884.

White, M.M., Siri, A., Lichtwardt, R.W. 2006. Trichomycete insect symbionts in Great Smoky Mountains National Park and vicinity. Mycologia, 98(2): 333-352. doi: 10.1080/15572536.2006.11832705

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wilke, C.O. 2020. Introduction to Cowplot, R Foundation for Statistical Computing, 15 Dec. 2020, cran.r-project.org/web/packages/cowplot/vignettes/introduction.html.

Yeates, D.K., Meusemann, K., Trautwein, M., Wiegmann, B., Zwick, A. 2016. Power, resolution and bias: recent advances in insect phylogeny driven by the genomic revolution. Current Opinion in Insect Science, 13: 16–23.

Yin, Y., Mao, X., Yang, J.C., Chen, X., Mao, F., Xu, Y. 2012. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res., 40(Web Server issue): W445-51.

Young, T.W.K. 1999. Taxonomy of Kickxellaceae and Radiation of the Asexual Apparatus. Kew Bulletin, 54(3): 651-661. https://www.jstor.org/stable/4110861

Zhang, C., Rabiee, M., Sayyari, E., Mirarab, S. 2018. ASTRAL-III: Polynomial Time Species Tree Reconstruction from Partially Resolved Gene Trees. BMC Bioinformatics, 19(S6): 153. doi:10.1186/s12859-018-2129-y.

Zhang, F., Ding, Y., Zhu, C.-D., Zhou, X., Orr, M.C., Scheu, S., Luan, Y.-X. 2019. Phylogenomics from low-coverage whole-genome sequencing. Methods in Ecology and Evolution, 10: 507–517. doi: 10.1111/2041-210X.13145

Zhang, Z., Phillips, R.P., Zhao, W., Yuan, Y., Liu, Q., Yin, H. 2018b. Mycelia-derived C contributes more to nitrogen cycling than root-derived C in ectomycorrhizal alpine forests. Functional Ecology, 33(2): 346–359. https://doi.org/10.1111/1365-2435.13236

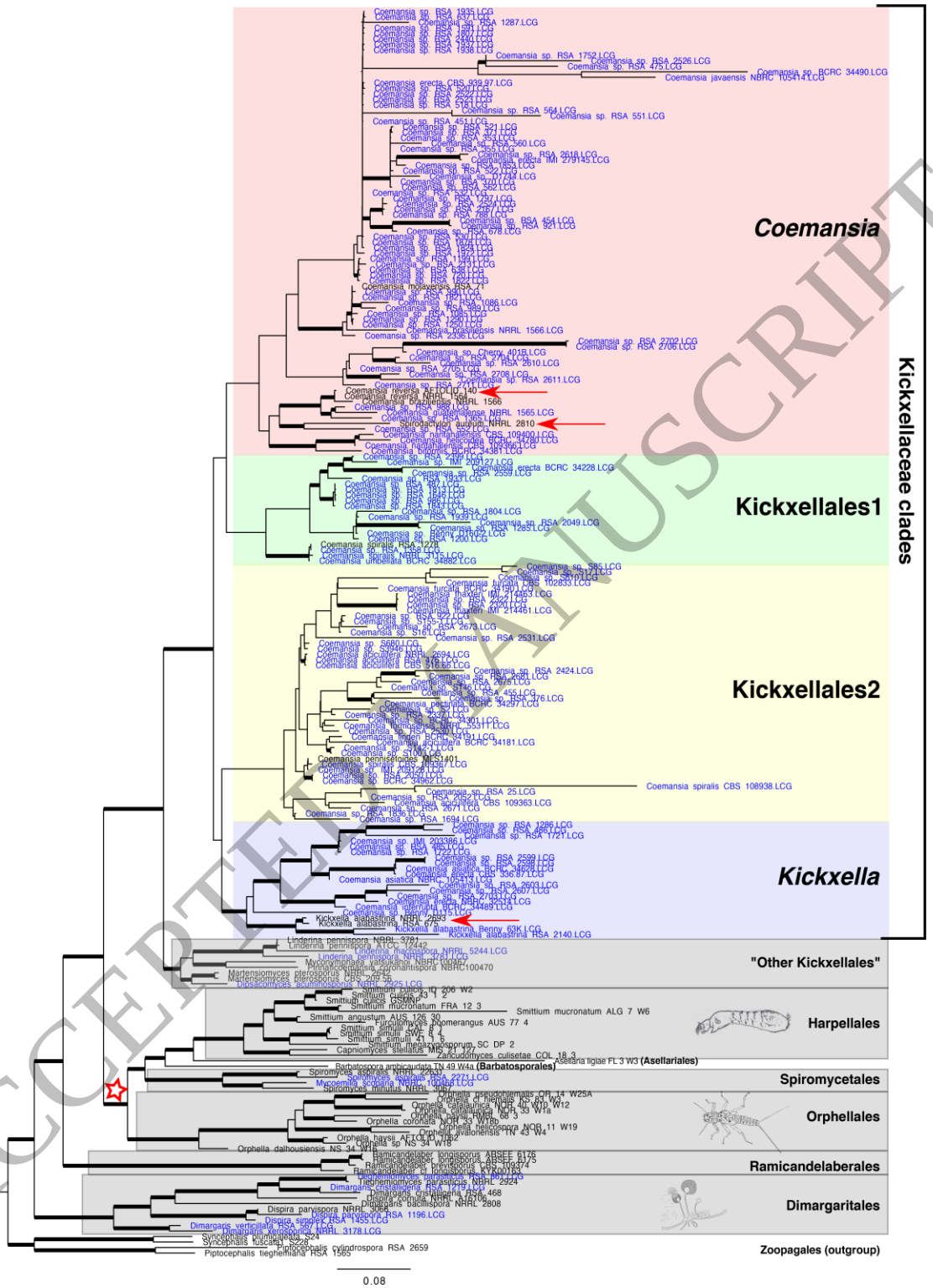

Figure 1
165x221 mm (1.3 x DPI)

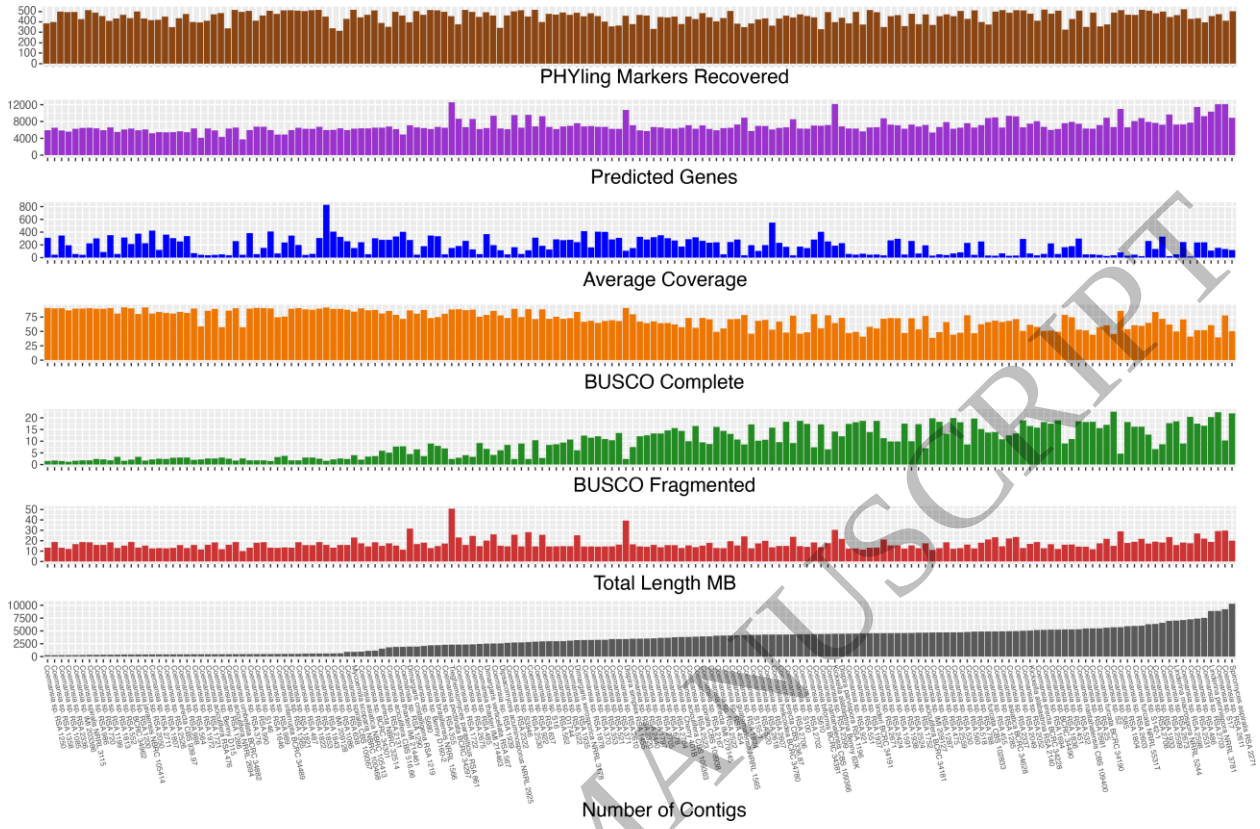

Figure 2

165x108 mm (1.3 x DPI)

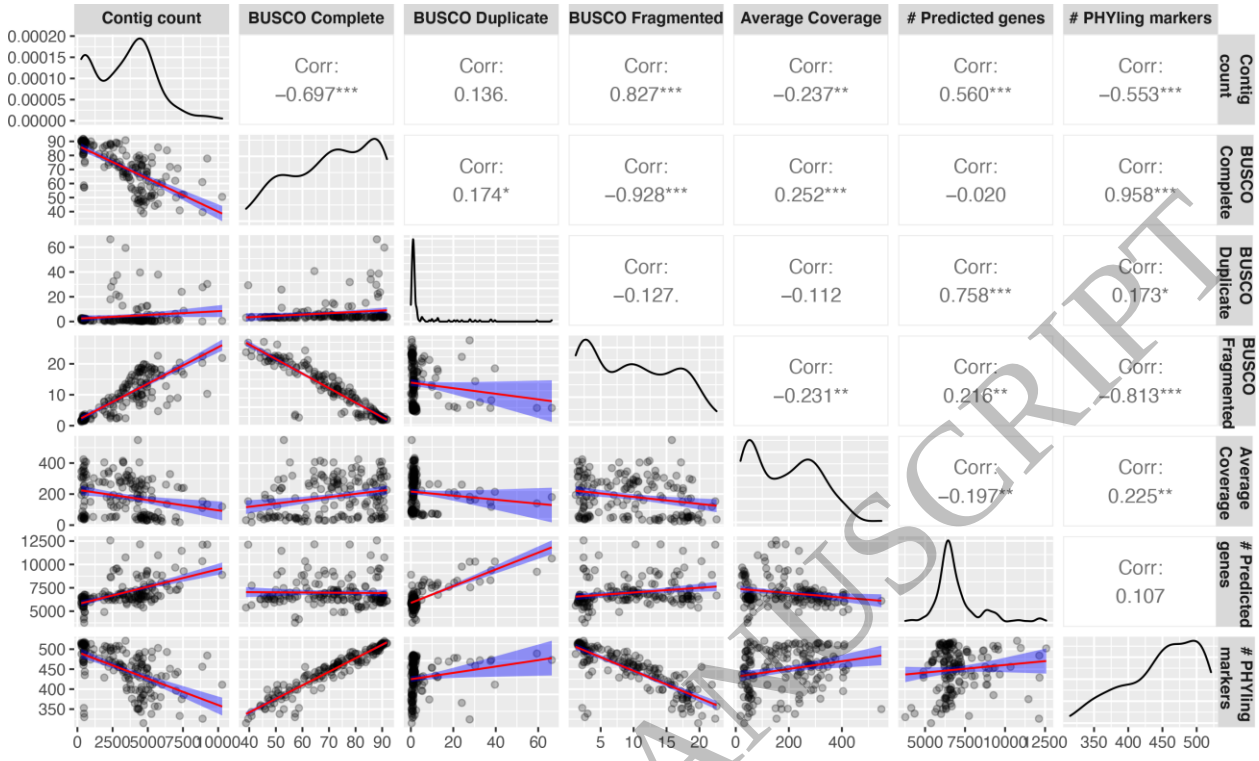

Figure 3
165x98 mm (1.3 x DPI)

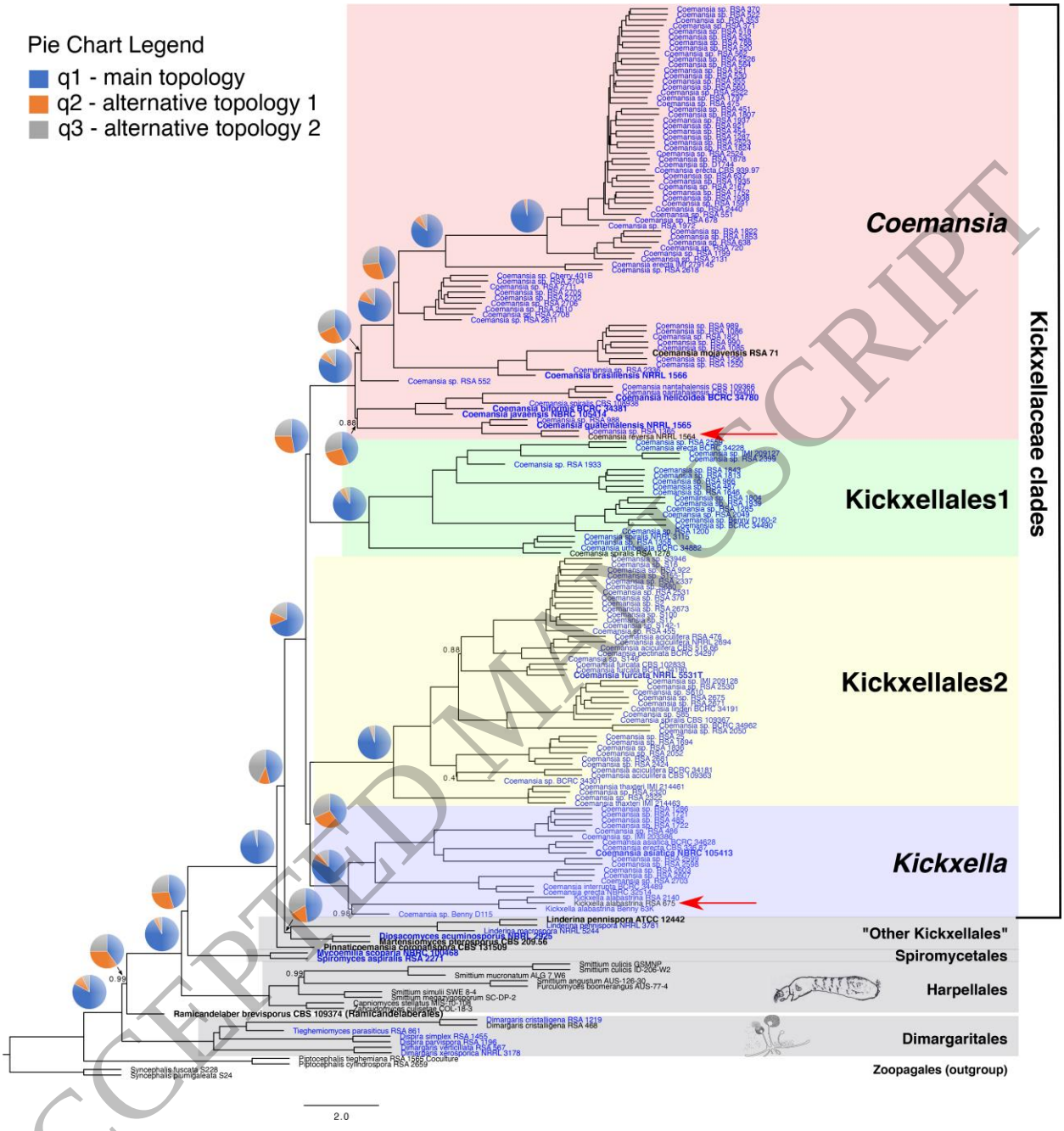

Figure 4
165x174 mm (1.3 x DPI)

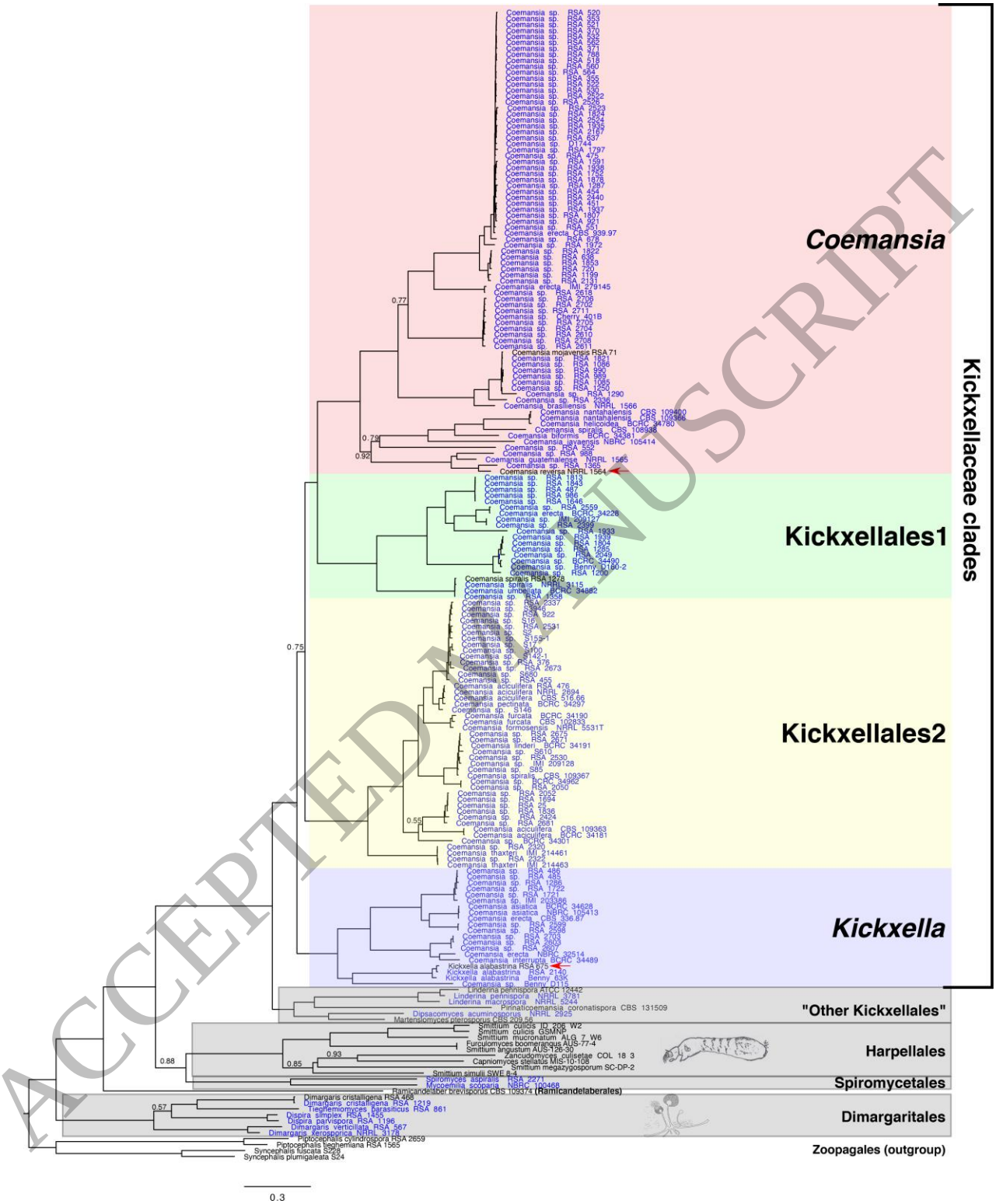

Figure 5
165x204 mm (1.3 x DPI)

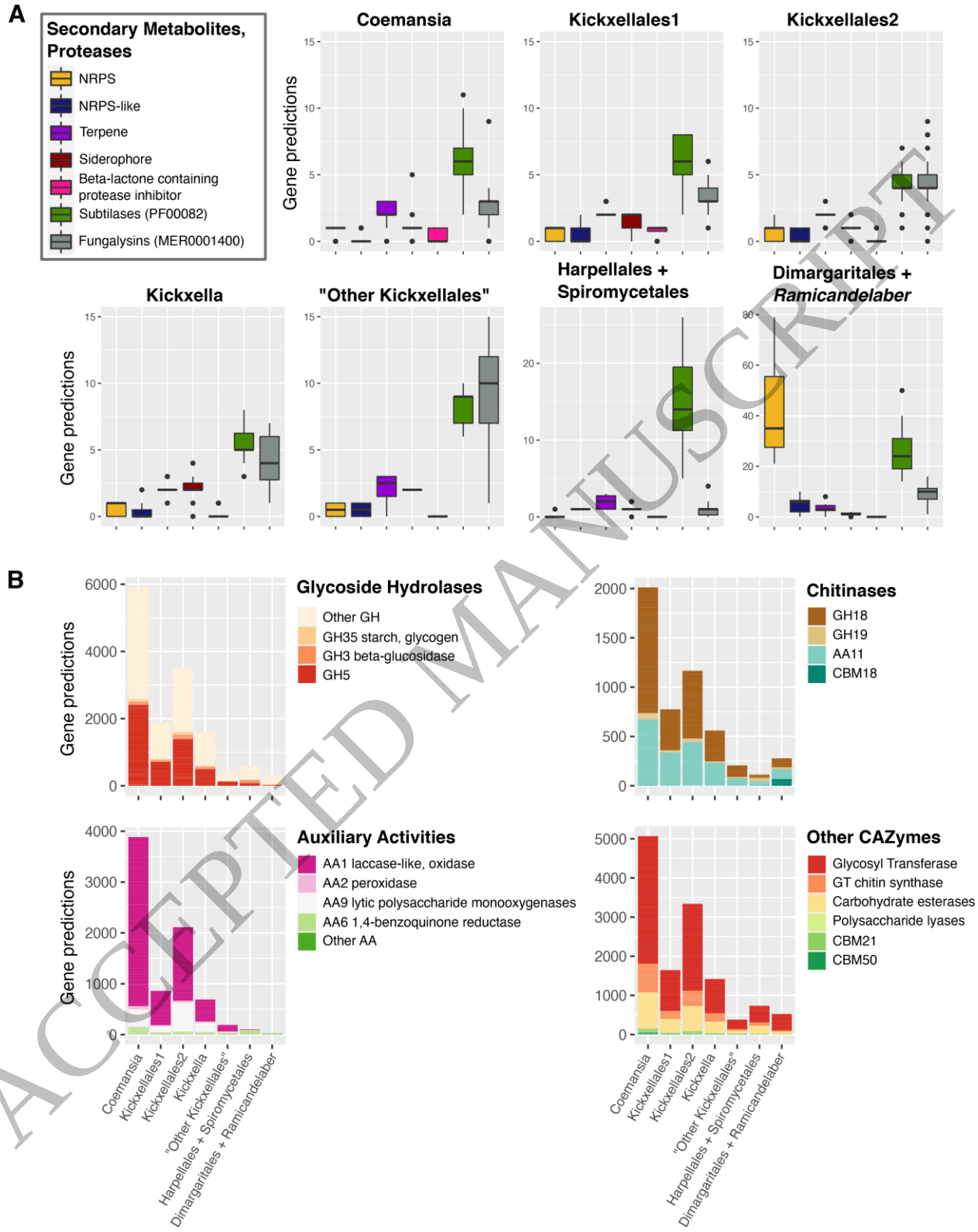

Figure 6
165x209 mm (1.3 x DPI)

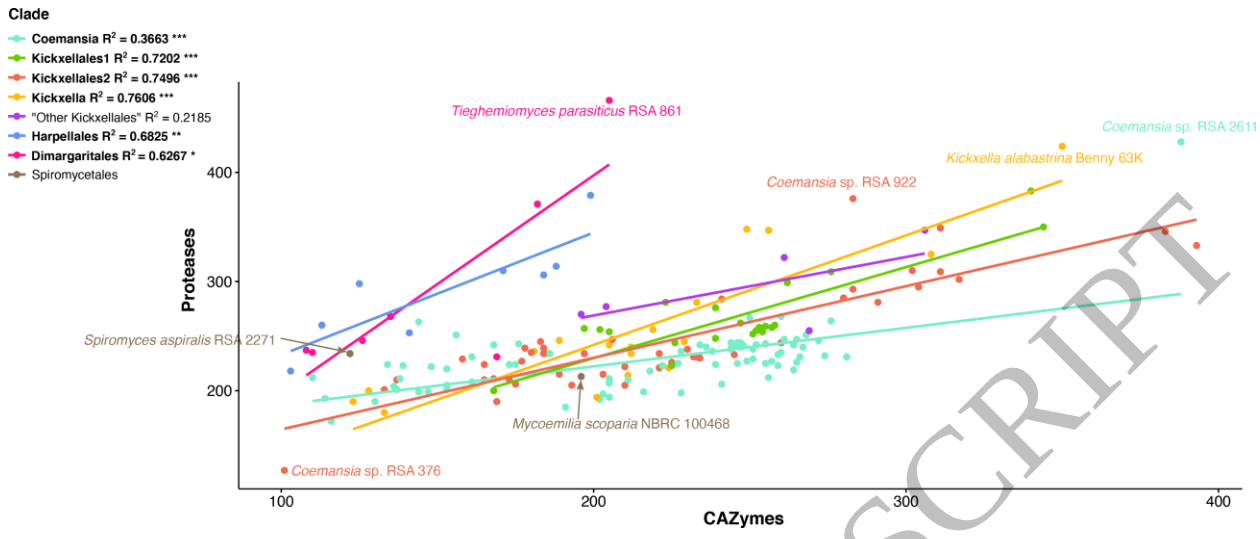

Figure 7
165x69 mm (1.3 x DPI)