Title

Convergent reductive evolution and host adaptation in Mycoavidus bacterial endosymbionts of Mortierellaceae fungi

Authors

Amses, Kevin

Desiró, Alessandro

Bryson, Abigail

et al.

Publication Date

2023-12-01

DOI

10.1016/j.fgb.2023.103838

Fungal Genetics and Biology 169 (2023) 103838

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Convergent reductive evolution and host adaptation in *Mycoavidus* bacterial endosymbionts of Mortierellaceae fungi

Kevin Amses [a], Alessandro Desiró [b], Abigail Bryson [b], Igor Grigoriev [c,d], Stephen Mondo [c,e], Anna Lipzen [c], Kurt LaButti [c], Robert Riley [c], Vasanth Singan [c], Paris Salazar-Hamm [a], Jason King [e], Elizabeth Ballou [f], Teresa Pawlowska [g], Rasheed Adeleke [h,i], Gregory Bonito [b], Jessie Uehling [a,*]

[a] Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR 97333, USA
[b] Department of Plant Soil and Microbial Sciences, Michigan State University, East Lansing MI 48824, USA
[c] United States Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
[d] Department of Plant and Microbial Biology, University of California Berkeley, Berkeley, CA 94720, USA
[e] Department of Agricultural Biology, Colorado State University, Fort Collins, CO, USA
[f] School of Biosciences, University of Sheffield, Western Bank S10 2TN, UK
[g] MRC Centre for Medical Mycology, University of Exeter, Exeter EX4 4QD, UK
[h] School of Integrative Plant Science, Cornell University, Ithaca, NY 14853-5904, USA
[i] Unit for Environmental Sciences and Management, North-West University, Potchefstroom, Private bag X6001, 2520, South Africa

ARTICLE INFO

Keywords:
Host-specificity
Fungal bacterial interactions
Secondary metabolism
Gene loss

ABSTRACT

Intimate associations between fungi and intracellular bacterial endosymbionts are becoming increasingly well understood. Phylogenetic analyses demonstrate that bacterial endosymbionts of Mucoromycota fungi are related either to free-living *Burkholderia* or *Mollicutes* species. The so-called *Burkholderia*-related endosymbionts or BRE comprise *Mycoavidus, Mycetohabitans* and *Candidatus* Glomeribacter gigasporarum. These endosymbionts are marked by genome contraction thought to be associated with intracellular selection. However, the conclusions drawn thus far are based on a very small subset of endosymbiont genomes, and the mechanisms leading to genome streamlining are not well understood. The purpose of this study was to better understand how intracellular existence shapes *Mycoavidus* and BRE functionally at the genome level. To this end we generated and analyzed 14 novel draft genomes for *Mycoavidus* living within the hyphae of Mortierellomycotina fungi. We found that our novel *Mycoavidus* genomes were significantly reduced compared to free-living Burkholderiales relatives. Using a genome-scale phylogenetic approach including the novel and available existing genomes of *Mycoavidus*, we show that the genus is an assemblage composed of two independently derived lineages including three well supported clades of *Mycoavidus*. Using a comparative genomic approach, we shed light on the functional implications of genome reduction, documenting shared and unique gene loss patterns between the three *Mycoavidus* clades. We found that many endosymbiont isolates demonstrate patterns of vertical transmission and host-specificity, but others are present in phylogenetically disparate hosts. We discuss how reductive evolution and host specificity reflect convergent adaptation to the intrahyphal selective landscape, and commonalities of eukaryotic endosymbiont genome evolution.

1. Introduction

Endosymbiotic bacteria have some of the smallest genomes of self-replicating organisms (McCutcheon et al., 2019; Wernegreen, 2015). Genome streamlining is a hallmark of eukaryotic endosymbiosis thought to be caused by unique selective pressures inside of cells. During streamlining, genes for cellular replication, host interactions and amino acid biosynthesis are either retained, degraded, or entirely lost in lineage specific patterns (Chong et al., 2019; Uehling et al., 2017; Wernegreen, 2002). Much of our understanding of endosymbiotic evolution comes from insect endosymbionts. In these systems, endosymbionts with small effective population sizes are limited to host resources and can

* Corresponding author.
E-mail address: Jessie.Uehling@OregonState.edu (J. Uehling).

https://doi.org/10.1016/j.fgb.2023.103838
Received 13 April 2023; Received in revised form 6 September 2023; Accepted 11 September 2023
Available online 15 September 2023

undergo extreme bottlenecks during transmission between individuals (Wernegreen, 2015). When these small populations are strictly vertically transmitted or isolated, endosymbiont genomes respond more rapidly to genetic drift or the intracellular selective pressures faster than bacteria in larger populations (Moran, 1996; Wernegreen and Moran, 1999; Woolfit and Bromham, 2003). In some cases, mutational biases and rapid evolutionary rates work together to pseudogenize genes leading to further and faster gene loss (Bennett and Moran, 2013; Waneka et al., 2021). Together these phenomena are thought to account for the observations that many endosymbiont genomes are physically much smaller than their free-living bacterial relatives and differ in having an altered functional composition of genes. In addition to insects, examples of streamlined endosymbionts include intracellular parasites, organelles, plastids, and endosymbionts of nematodes and fungi (McCutcheon et al., 2019). Because genome streamlining directly reflects the functional basis of symbioses, endosymbiont genomes offer a unique lens into how endosymbionts retain adaptive genomic content in response to intracellular selective pressures.

Mucoromycota fungi are ecologically diverse, speciose, and notorious hosts of Mycoplasma and Burkholderiales-related endosymbionts (Spatafora et al. 2016; Uehling et al. 2023); Bianciotto et al., 2003; Bonfante and Desirò, 2017; Itabangi et al., 2022; Lackner et al., 2011; Mondo et al., 2012; Ohshima et al., 2016). These fungi are coenocytic and have diverse metabolic potentials reflective of the physical environments in which each fungal lineage is found. For example, Mortierellomycotina fungi are saprotrophic and plant-associated with genomes enabled to utilize simple plant-based sugars and to produce high amounts of lipids (Liao et al., 2019; Uehling et al., 2017; Zhao et al., 2022). In contrast, Glomeromycotina fungi are obligate plant biotrophic symbionts, with genomes enriched in transporters and metabolic machinery for plant sugars but lacking fatty acid synthetase genes (Luginbuehl et al., 2017). Lastly, Mucoromycotina fungi are largely saprotrophic and associated with decaying plant material and have genomes containing enzymes for the breakdown of simple organic carbon sources (Gryganskyi et al., 2018). Each of the three Mucoromycota subphyla may host endosymbiotic bacteria that are specialized in obtaining intracellular resources from their hosts, reflective of their intracellular status (Bonfante and Desirò, 2017; Gryganskyi et al., 2018; Mondo et al., 2017; Uehling et al., 2017); Ohshima et al., 2016; Partida-Martinez et al., 2007; Sato et al., 2010). These bacteria enter fungal cells via mechanisms that are currently being studied, and the degree of their horizontal transmission is largely unknown. While several distinct bacterial taxa have been documented to co-reside in the same fungal species, and even isolate (Hoffman and Arnold 2010) the most common observations of fungi with endosymbionts include a single bacterial species inhabiting a fungal isolate. Among the best studied fungal-endosymbiont systems are *Candidatus* Glomeribacter gigasporarum (CaGg) with fungal taxa in the Glomeromycotina, *Mycetohabitans* spp. with *Rhizopus* spp. in the Mucoromycotina, and *Mycoavidus cysteinexigens* isolates with fungal taxa in the Mortierellomycotina. Much remains to be learned about functional symbiotic dynamics between fungi and bacterial endosymbionts, but a significant symbiotic benefit to host fungi is thought to be the production of endosymbiont secondary metabolites (Büttner et al., 2021; Richter Ingrid et al., 2022; Scherlach et al., 2006).

One well studied BRE endosymbiont is *Mycoavidus cysteinexigens*, which is best known for interactions with *Linnemannia elongata* (=*Mortierella elongata*) and recently documented in species of *Actinomortierella*, *Linnemannia*, *Lunasporangiospora*, *Mortierella* and *Podila* (Büttner et al., 2021; Herlambang Afri et al., 2022; Takashima et al., 2018; Telagathoti et al., 2021). *Mycoavidus* is currently thought to be monophyletic, and comprises three well supported clades known as A, B, and C, based on available 16S ribosomal rDNA sequences (Sharmin et al. 2018; Okrasińska et al. 2021; Takashima et al. 2018; Uehling et al. 2023). While comparative analyses utilizing available *Mycoavidus* genomes suggests ~ 80% of their genes are shared and have high homology, unique gene content and significant rearrangements between individuals have been documented (Guo et al., 2020). There are five publicly available *Mycoavidus* genomes that range in size from 1.9 to 2.8 Mb, in stark contrast to their free-living relatives that have genomes ranging from 5 to 12 Mb (Büttner et al., 2021; Guo et al., 2020; Sharmin et al., 2018; Uehling et al., 2017). Genome contraction in *Mycoavidus* compared to free-living relatives is thought to reflect adaptation to fungal host resources, in particular amino acids and fatty acids (Guo et al., 2020; Ohshima et al., 2016; Uehling et al., 2017). For example, *Mycoavidus cysteinexigens* AG77 lacks complete pathways for glycolysis and some amino acid biosynthesis, and instead relies on host fungal amino acids and fatty acids as evidenced by bacterial transporter capacity and accumulation of these products in endosymbiont free, cleared fungal isolates (Li et al., 2017; Uehling et al., 2017). Together, these observations have led to the understanding that *Mycoavidus* are obligate, host-dependent, intracellular endosymbionts with streamlined genomes that reflect metabolic adaptation to intrahyphal life. So far, high specificity in associations of *Mycoavidus* and Mortierellaceae fungi have been observed, and it is hypothesized that this is due to primarily vertical transmission between generations in this system (Guo et al., 2020; Li et al., 2017; Sharmin et al., 2018; Uehling et al., 2017). However, because *Mycoavidus* genomes are scarce, host specificity, ubiquity of genome streamlining, and functional categorization of gene loss in *Mycoavidus* genomes has yet to be quantified on a broad scale.

The influence of bacterial endosymbionts, including *Mycoavidus,* on fungal hosts has been primarily gleaned from fungal physiology data comparing behavior of isogenic fungal isolates with and without endosymbionts (Büttner et al., 2021; Itabangi et al., 2022; Li et al., 2017; Mondo et al., 2017; Richter Ingrid et al., 2022; Uehling et al., 2017). When *Mycoavidus* is present in fungal hyphae, hosts have altered sexual spore production rates (Takashima et al., 2020) and unique volatile production patterns (Misztal et al., 2018). In addition, *Linnemannia* grows faster without *Mycoavidus* (Uehling et al., 2017), reflecting altered fungal metabolism of amino acids, fatty acids, and sugars (Li et al., 2017). Because endosymbiont genome reduction reflects dependence on host derived resources, we hypothesized that newly sequenced *Mycoavidus* isolates would exhibit convergent loss in pathways reflecting similar host metabolic adaptation.

In the present study, we evaluated whether *Mycoavidus* isolates from geographically diverse Mortierellaceae hosts share gene loss patterns reflective of host adaptation based on whole genome sequence data. We then used the resulting genomic contigs to evaluate diversity and molecular evolution of *Mycoavidus* isolates among extant clades of BRE, and to assess fungal host specificity. We found that the three distinct *Mycoavidus* clades A-C are well supported with phylogenomic data and differ in biology including host-specificity. In addition, we found that *Candidatus* Glomeribacter gigasporarum is nested between *Mycoavidus* clades A and B/C, rendering *Mycoavidus* in need of taxonomic revision. Further we observed that *Mycoavidus* genomes reflect convergent gene loss in amino acid pathways, yet the missing genes are not always the same. We discuss convergent intracellular selective pressures on endosymbiont genomes, and how these have shaped the evolutionary history of *Mycoavidus* and relatives.

2. Materials and methods

2.1. Isolate source, culture conditions and DNA extraction

Fungal strains were isolated from soils using previously described techniques (Desirò et al., 2023) (Table 1). Fungal isolates were transferred to 1% (w/v) Malt Extract Agar (MEA, Difco, Detroit MI) plates supplemented with 50 μg/mL kanamycin, chloramphenicol, and ampicillin. After 48 h of incubation at room temperature, and an additional 48 h of subculture on the same media, fungal isolates were transferred to MEA plates without antibiotics covered with a sterile cellophane membrane. After at least 14 days of growth at room temperature, fungal

Table 1
Table summarizing fungal hosts, NCBI endosymbiont genome accessions, geographic isolation origin, *Mycoavidus* BRE clade status, and genome assembly metadata for genomes generated in this study and used in analyses presented here.

Taxon	strain	Genbank accession	genome size (bp)	contigs	host	*Mycoavidus* BRE clade	isolate origin	BUSCO completeness
Mycoavidus cysteinexigens	AV005	JASSUI000000000	3.2	3	*Actinomortierella capitata*	C	Puerto Rico	89.10%
Mycoavidus cysteinexigens	AD051	JASSUH000000000	2.9	5	*Podila minutissima*	C	Michigan, USA	89.68%
Mycoavidus cysteinexigens	AD073	JASSUG000000000	2.6	16	*Linnemannia elongata*	A	Michigan, USA	89.68%
Mycoavidus cysteinexigens	AD266	JASSUJ000000000	2.4	1	*Mortierella alpina*	C	Oregon, USA	89.83%
Mycoavidus cysteinexigens	AD086	JASSUL000000000	2.2	28	*Podila humilis*	A	Pennsylvania, USA	86.63%
Mycoavidus cysteinexigens	AD058	JASSUD000000000	2.8	134	*Podila epicladia*	A	Michigan, USA	86.19%
Mycoavidus cysteinexigens	AM1000	JASSUF000000000	2.4	235	*Podila clonocystis*	B	Illinois, USa	82.99%
Mycoavidus cysteinexigens	TTC192	JASSUE000000000.	2.3	42	*Podila verticillata*	A	North Carolina, USA	90.41%
Mycoavidus cysteinexigens	AD045	JASSUC000000000	2.3	79	*Linnemannia gamsii*	A	Michigan, USA	89.53%
Mycoavidus cysteinexigens	AD022	JASSUM000000000	2.6	151	*Linnemannia elongata*	A	Utah, USA	89.83%
Mycoavidus cysteinexigens	NVP60	JASSUB000000000	2.4	95	*Linnemannia gamsii*	A	Michigan, USA	89.24%
Mycoavidus cysteinexigens	AM980	JASSUN000000000	2.6	132	*Linnemannia elongata*	A	Illinois, USa	89.83%
Mycoavidus cysteinexigens	AD003	JASSUO000000000	2.1	87	*Podila humilis*	A	New Zealand	90.41%
Mycoavidus cysteinexigens	REB-025A	JASSUK000000000	2.5	56	*Podila minutissima*	A	North Carolina, USA	86.48%

mycelium was harvested, manually homogenized with a pestle, and genomic DNA was extracted using a modified version of the CTAB protocol (Uehling et al., 2017; Desirò et al., 2023). Briefly, homogenized tissues were incubated for 60 min at 65 °C in CTAB buffer supplemented with 1% (w/v) beta-mercaptoethanol and 100 μg proteinase K. Following incubation, nucleic acids were extracted in two subsequent rounds of polar phase separation using 24:1 chloroform isoamyl alcohol. Nucleic acids were precipitated with isopropyl alcohol and the resulting pellets washed with 70%(w/v) and then 100% ethanol. Cleaned pellets were solubilized in sterile molecular grade water.

2.2. Genomic DNA isolation, sequencing, and assembly

Genomic DNA was isolated using the CTAB protocol described above. The combined genomic DNA of each host-symbiont pair was sequenced using either the PacBio or Illumina platforms (Table 1). For genomes sequenced using PacBio technology, ~5 μg of genomic DNA was sheared to > 10 Kb using Covaris g-Tubes (Covaris Biosciences, Woburn MA) and size selected for fragments > 10 Kb using AMPure Beads (Thermo-Fisher, Waltham MA). The sheared DNA was then treated with exonuclease to remove single-stranded ends and DNA was repaired using DNA damage repair mix followed by end repair and ligation of blunt adapters with SMRTbell Template Prep Kit 1.0 (Pacific Biosciences, Menlo Park CA). The library was then purified with AMPure PB beads. Size-selection targeting fragments>6 Kb was performed with the BluePippin system (Sage Science) for a significant 2 Kb fraction was observed in the purified library. PacBio Sequencing primer was then annealed to the SMRTbell template library and Version P6 sequencing polymerase was bound to them. The prepared SMRTbell template libraries were sequenced on Pacific Biosciences RSII or SEQUEL sequencers using Version C4 chemistry and 1x240 (RSII) and 1x300 or 1x600 (SEQUEL) sequencing movi e run times. For genomes sequenced using Illumina technology, 500 bp-plate-based DNA library preparation for Illumina sequencing was performed on the PerkinElmer Sciclone NGS robotic liquid handling system (PerkinElmer, Tempe AZ) using Kapa Biosystems library preparation kit (Kapa Biosciences, Wilmington MA). A total of 200 ng of sample DNA was sheared to 600 bp using a Covaris LE220 focused-ultrasonicator. The sheared DNA fragments were size selected by double-SPRI and then the selected fragments were end-repaired, A-tailed, and ligated with Illumina compatible sequencing adaptors from IDT containing a unique molecular index barcode for each sample library. The prepared libraries were quantified using KAPA Biosystems' next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument (Roche, Basel Switzerland). The quantified libraries were then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v4 (Illumina, San Diego). Sequencing of the flowcell was performed on the Illumina HiSeq2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a 2x150 indexed run recipe.

PacBio subread sequence data was processed with the JGI QC pipeline to remove artifacts. Filtered subread data was assembled together to generate an initial assembly using Falcon v.0.4.2–1.8.8, Celera v.1.8, and Flye.v.2.4–2.5 (Denisov et al., 2008; Freire et al., 2022; Kronenberg et al., 2018). Assembled endosymbionts were identified from the initial assembly using BLAST to NCBI nt, GC, Coverage, and Tetramer Nucleotide Frequency (TNF) PCA analysis, separated from the initial assembly, and improved separately using read recruitment with BBtools version 38.76 [bbduk.sh k = 31 mm = f mkf = 0.05] and subsequent reassembly of matching reads with Flye version 2.5 [–pacbio-corr –asm-coverage 50]. In some cases, multiple rounds of recruitment and gap closure with finisherSC version 2.1 (Lam et al., 2014) was required for improvement. Final assemblies were polished with Arrow version SMRTLink v6.0.0.47841 or gcpp SMRTLINK v8.0.0.80529 [–algorithm arrow] (https://github.com/PacificBiosciences/GenomicConsensus, https://www.pacb.com/support/software-downloads).

Illumina short read libraries composed of both host and symbiont sequences were assembled with SPAdes v3.15.3 (Prjibelski et al., 2020). Illumina-sequenced draft metagenomic assemblies were filtered using a combination of metabat2 v2 (Kang et al., 2019) and SCGid (Amses et al., 2020). Bacterial bins resulting from metabat2 were identified either by the extraction and analysis of 16S rDNA sequences detected with

barrnap v.09 (Seemann, 2015) or, when 16S rDNA sequences were not present, the taxonomy of top BLAST hits in the NCBI nucleotide database. Filtering with metabat2 and SCGid produced filtered bacterial draft genomes of similar sizes and completeness. The final filtered draft assemblies contain merged contigs that were either present in both the metabat2 and SCGid drafts or present in only one of these if the taxonomy of the most significant BLAST hit in the NCBI nt database was bacterial. rDNA-containing contigs identified from genomic data by barrnap were also manually included into final filtered assemblies. Gene sequences for 16S rDNA phylogenies (Fig. 1) were collected from the NCBI GenBank or extracted from the endosymbiont genomes assembled here using barrnapp (Table 2). All assembled genomes are deposited in NCBI GenBank (Table 1).

2.3. Genome annotation and analyses

Bacterial genomes were extracted from metagenomic backgrounds by the methods described above and annotated with Prokka v1.14.6 (Seemann et al. 2014) using translation tab. 11. While some fungal strains are known to harbor multiple bacteria (e.g. AD073) we focused on *Mycoavidus* genomes. Summarized annotation results are shown along with extracted assembly statistics in Table 1. Fungal genomes were annotated using funannotate v 1.8.14 (https://github.com/nextgenusfs/funannotate). In addition to the 14 *Mycoavidus* genomes generated herein, 56 additional BRE and non-BRE Burkholderiaceae genomes were selected for inclusion in phylogenetic and comparative analyses to maximize representative coverage of the major clades in Burkholderiaceae (Table S1). To evaluate the role of genome completeness in our analyses, we included the genomes of *Mycoavidus cysteinexigens* AG77 in two formats, an earlier fragmented Illumina based format (Patric 224135.3), and a later complete circular bacterial genome based on the inclusion of Pacbio sequencing reads (JAPELI000000000).

2.4. Phylogenetics

To create phylogenomic trees we utilized custom scripts available at https://github.com/Michigan-Mycology/Chytrid-Phylogenomics. For each endosymbiont genome, the predicted proteome was searched against the BUSCO *burkholderiales_odb10* database, which is composed of 688 orthologs, using the hmmsearch function included in *hmmer* v3.3.2 (Eddy 2015). To avoid introduction of paralogous sequences into concatenated phylogenetic analyses, all top hits (score_cutoffs < 1e-3) for each protein were retained, resulting in zero to many sequences from each genome. For each top hit, alignments were generated with hmmalign (Eddy 2015), trimmed with trimal (Capella-Gutiérrez et al., 2009), and highly gapped sequences were removed at a threshold of > 0.75. Gene trees were created with fasttree v2.1.11 (Price et al., 2010) and were curated automatically, using an integrative approach that yielded individual locus alignments including one sequence per taxon and filtered as follows. Gene trees were traversed, and it was determined whether all tips (top hit sequences from a genome) were monophyletic. If they were monophyletic, top hits were retained. If tips of a gene tree were not monophyletic, scores of each protein in the gene tree (i.e., tips) were compared to see if lack of monophyly was the result of one or more particularly low-scoring clusters of tips (<= 70% of score of highest-scoring tip for that genome). Low-scoring tips were removed, and tips were checked again for monophyly. If removal of low-scoring tips led to monophyly of the taxon, the highest-scoring tip sequence among the high-scoring cluster was taken. If monophyly did not result from removal of the low-scoring tips, they were permanently removed. All tips with hits in the higher-scoring bin were retained despite their polyphyly or paraphyly nature. In a second round of score-based filtering, we removed tips with scores that were lower than 1.5 standard deviations from the mean tip score calculated from all tips in each gene tree. In this way, we removed low-scoring tips from the tree that received a score higher than 70% of the taxon-specific maximum but were low-scoring relative to the entire tree. Raw marker occupancy

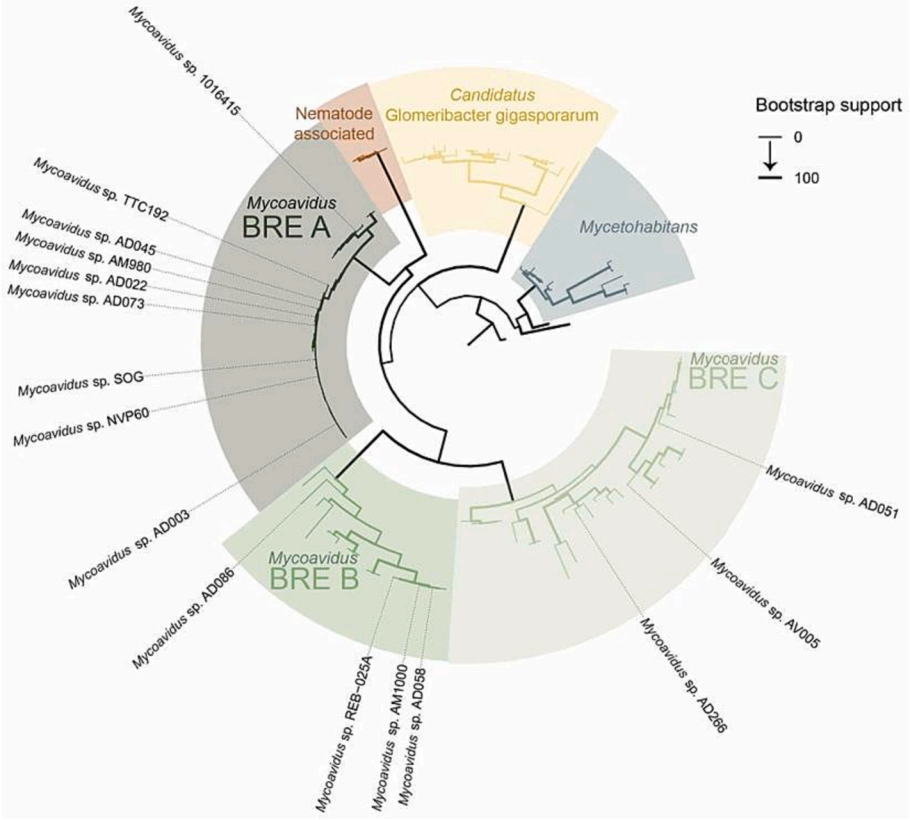

Fig. 1. Phylogenetic tree of Burkholderiaceae including novel *Mycoavidus* isolates based on 16S rDNA sequences supporting a monophyletic *Mycoavidus*. Major groups of *Mycoavidus* BRE are highlighted by colored boxes and corresponding labels. Tips corresponding to novel genome data generated in this study are labeled while all publicly available 16S sequences are not (see Table S2 for a full list of included sequences). Newly sequenced *Mycoavidus* are distributed across previously described lineages.

Table 2
Table containing NCBI accessions for publicly available 16S rDNA sequences used in this study and analyses presented here.

Accession	Taxon
AF043302	*Burkholderia ambifaria*
U96927	*Burkholderia cepacia*
AF148556	*Burkholderia cepacia*
U96927	*Burkholderia cepacia*
AM747629	*Burkholderia diffusa*
AB680484	*Burkholderia gladioli*
EU024168	*Burkholderia gladioli*
GU936678	*Burkholderia gladioli*
U96931	*Burkholderia glumae*
AF110188	*Burkholderia mallei*
AM747632	*Burkholderia metallica*
U96930	*Burkholderia pyrrocinia*
AM747631	*Burkholderia seminalis*
MT002691	*Burkholderia* sp.
MT002716	*Burkholderia* sp.
MW055707	*Burkholderia* sp.
MW055867	*Burkholderia* sp.
MW080027	*Burkholderia* sp.
MW080031	*Burkholderia* sp.
KU899555	*Burkholderia* sp. *nematode symbiont*
KU899558	*Burkholderia* sp. *nematode symbiont*
KT735077	*Burkholderia* sp. *nematode symbiont*
LK023502	*Burkholderia stagnalis*
U91838	*Burkholderia thailandensis*
AF097534	*Burkholderia vietnamiensis*
EU723243	*Burkholderia xenovorans*
X89727	*Candidatus* Glomeribacter gigasporarum
AJ251636	*Candidatus* Glomeribacter gigasporarum
KF378650	*Candidatus* Glomeribacter gigasporarum
KF378649	*Candidatus* Glomeribacter gigasporarum
AJ251633	*Candidatus* Glomeribacter gigasporarum
AM889130	*Candidatus* Glomeribacter gigasporarum
AM889132	*Candidatus* Glomeribacter gigasporarum
AJ251635	*Candidatus* Glomeribacter gigasporarum
KF378652	*Candidatus* Glomeribacter gigasporarum
KF378651	*Candidatus* Glomeribacter gigasporarum
AJ251634	*Candidatus* Glomeribacter gigasporarum
AM889129	*Candidatus* Glomeribacter gigasporarum
AM889128	*Candidatus* Glomeribacter gigasporarum
KF378648	*Candidatus* Glomeribacter gigasporarum
AM889131	*Candidatus* Glomeribacter gigasporarum
EU625665	*Candidatus* Glomeribacter gigasporarum
FN252291	*Candidatus* Glomeribacter gigasporarum
AJ251633	*Candidatus* Glomeribacter gigasporarum
AJ251634	*Candidatus* Glomeribacter gigasporarum
AJ251635	*Candidatus* Glomeribacter gigasporarum
KT944310	*Candidatus* Pandoraea novymonadis
JN810871	*Candidatus* Vallotia
JN810866	*Candidatus* Vallotia
JN810867	*Candidatus* Vallotia
JN810865	*Candidatus* Vallotia
JN810874	*Candidatus* Vallotia
KT735068	*Candidatus* Xiphinematincola pachtaicus nematode symbiont
DQ256728	*Chitinimonas koreensis*
AY323827	*Chitinimonas taiwanensis*
AY281146	*Collimonas arenae*
AJ310394	*Collimonas fungivorans*
AY281137	*Collimonas pratensis*
AF191737	*Cupriavidus necator*
AB121221	*Cupriavidus pinatubonensis*
Y08845	*Janthinobacterium agaricidamnosum*
Y08846	*Janthinobacterium lividum*
AB366174	*Limnobacter litoralis*
AJ289885	*Limnobacter thiooxidans*
AM420302	*Mycetohabitans endofungorum*
AM420302	*Mycetohabitans endofungorum*
HQ005412	*Mycetohabitans endofungorum*
NR042393	*Mycetohabitans rhizoxinica*
HQ005411	*Mycetohabitans rhizoxinica*
HQ005408	*Mycetohabitans rhizoxinica*
AJ938142	*Mycetohabitans rhizoxinica*
MF383462	*Mycoavidus* clade A
MH760813	*Mycoavidus* clade A
MF383428	*Mycoavidus* clade A
MH760811	*Mycoavidus* clade A
MF383419	*Mycoavidus* clade A
KP772725	*Mycoavidus* clade A
KP772723	*Mycoavidus* clade A
MF383440	*Mycoavidus* clade A
MF383439	*Mycoavidus* clade A
MF383442	*Mycoavidus* clade A
MF383441	*Mycoavidus* clade A
MF383427	*Mycoavidus* clade A
MF383437	*Mycoavidus* clade A
MF383438	*Mycoavidus* clade A
MF383420	*Mycoavidus* clade A
AB558491	*Mycoavidus* clade A
KP772716	*Mycoavidus* clade A
AB558492	*Mycoavidus* clade A
LC005489	*Mycoavidus* clade A
MF383437	*Mycoavidus* clade A
MF383438	*Mycoavidus* clade A
MF383420	*Mycoavidus* clade A
MF383418	*Mycoavidus* clade A
LC005489	*Mycoavidus* clade A
NR149240	*Mycoavidus* clade A
MF383419	*Mycoavidus* clade A
MF383427	*Mycoavidus* clade A
MF383428	*Mycoavidus* clade A
MF383439	*Mycoavidus* clade A
MF383440	*Mycoavidus* clade A
MF383441	*Mycoavidus* clade A
MF383442	*Mycoavidus* clade A
MH760811	*Mycoavidus* clade A
MF383462	*Mycoavidus* clade A
MH760813	*Mycoavidus* clade A
MF383456	*Mycoavidus* clade B
MH760812	*Mycoavidus* clade B
AB558493	*Mycoavidus* clade B
MF383457	*Mycoavidus* clade B
MH760809	*Mycoavidus* clade B
MF383461	*Mycoavidus* clade B
MF383426	*Mycoavidus* clade B
MF383430	*Mycoavidus* clade B
MF383425	*Mycoavidus* clade B
MF383434	*Mycoavidus* clade B
MF383454	*Mycoavidus* clade B
MF383454	*Mycoavidus* clade B
MF383425	*Mycoavidus* clade B
MF383434	*Mycoavidus* clade B
MF383426	*Mycoavidus* clade B
MF383430	*Mycoavidus* clade B
MF383461	*Mycoavidus* clade B
MH760809	*Mycoavidus* clade B
MF383457	*Mycoavidus* clade B
MF383456	*Mycoavidus* clade B
MH760812	*Mycoavidus* clade B
MF383459	*Mycoavidus* clade C
MF383460	*Mycoavidus* clade C
MF383424	*Mycoavidus* clade C
MF383423	*Mycoavidus* clade C
MF383455	*Mycoavidus* clade C
MF383452	*Mycoavidus* clade C
MF383443	*Mycoavidus* clade C
MF383435	*Mycoavidus* clade C
MF383436	*Mycoavidus* clade C
MF383433	*Mycoavidus* clade C
MF383431	*Mycoavidus* clade C
MF383432	*Mycoavidus* clade C
MF383422	*Mycoavidus* clade C
MF383421	*Mycoavidus* clade C
MF383446	*Mycoavidus* clade C
MF383444	*Mycoavidus* clade C
MF383445	*Mycoavidus* clade C
MF383429	*Mycoavidus* clade C
MF383458	*Mycoavidus* clade C
MF383449	*Mycoavidus* clade C
MF383453	*Mycoavidus* clade C
MF383450	*Mycoavidus* clade C

(*continued on next page*)

Table 2 (*continued*)

Accession	Taxon
MH760810	*Mycoavidus* clade C
MF383451	*Mycoavidus* clade C
AB558494	*Mycoavidus* clade C
MF383447	*Mycoavidus* clade C
MF383448	*Mycoavidus* clade C
MF383450	*Mycoavidus* clade C
MF383451	*Mycoavidus* clade C
MF383453	*Mycoavidus* clade C
MH760810	*Mycoavidus* clade C
MF383421	*Mycoavidus* clade C
MF383446	*Mycoavidus* clade C
MF383422	*Mycoavidus* clade C
MF383447	*Mycoavidus* clade C
MF383448	*Mycoavidus* clade C
MF383449	*Mycoavidus* clade C
MF383458	*Mycoavidus* clade C
MF383429	*Mycoavidus* clade C
MF383444	*Mycoavidus* clade C
MF383445	*Mycoavidus* clade C
MF383431	*Mycoavidus* clade C
MF383432	*Mycoavidus* clade C
MF383433	*Mycoavidus* clade C
MF383435	*Mycoavidus* clade C
MF383436	*Mycoavidus* clade C
MF383443	*Mycoavidus* clade C
MF383452	*Mycoavidus* clade C
MF383455	*Mycoavidus* clade C
MF383423	*Mycoavidus* clade C
MF383424	*Mycoavidus* clade C
MF383459	*Mycoavidus* clade C
MF383460	*Mycoavidus* clade C
AF139173	*Pandoraea apista*
EF397578	*Pandoraea thiooxydans*
KC817488	*Paraburkholderia aspalathi*
AM489501	*Paraburkholderia bryophila*
EF139186	*Paraburkholderia caballeronis*
HQ698908	*Paraburkholderia dilworthii*
AF215705	*Paraburkholderia fungorum*
LN868266	*Paraburkholderia fungorum*
U96939	*Paraburkholderia graminis*
FJ796457	*Paraburkholderia humisilvae*
KF733462	*Paraburkholderia insulsa*
KJ601731	*Paraburkholderia jirisanensis*
KF155692	*Paraburkholderia monticola*
AY497470	*Paraburkholderia phytofirmans*
AB365791	*Paraburkholderia rhizosphaerae*
EU219865	*Paraburkholderia rhynchosiae*
EU035613	*Paraburkholderia sediminicola*
FJ772068	*Paraburkholderia solisilvae*
AJ302311	*Paraburkholderia tuberum*
NR136871	*Paraburkholderia ultramafica*
AB024310	*Paraburkholeria ururiensis*
AJ420332	*Paraburkhollderia tropicalis*
X92555	*Paucimonas lemoignei*
AJ879783	*Polynucleobacter asymbioticus*
AJ550672	*Polynucleobacter cosmopolitanus*
AM397067	*Polynucleobacter necessariu*
AY741342	*Ralstonia pickettii*
X67036	*Ralstonia solanacearum*
AF300324	*Ralstonia taiwanensis*
AY833061	*Wolbachia pipientis*

varied from 0.60 to 1.00 (mean = 0.94). After removing thirty-six low-occupancy markers, (<75% occupied) 652 markers remained and were used to conduct a concatenated phylogenetic analysis.

Individual gene alignments were concatenated and a Maximum Likelihood (ML) phylogenomic tree was computed in IQTree v1.6.9 (Mihn et al. 2020) with 100 nonparametric bootstraps under the Q.pfam + I + G4 model, selected as the best model by ModelFinder (Kalyaanamoorthy et al. 2017). To assess gene tree support for the best ML tree we utilized the gene tree concordance factor (gCF) implementation in IQTree2 as well as quartet support values as defined in ASTRAL-III (Zhang et al. 2018). Additionally, we implemented custom scripts (https://github.com/amsesk/EHB-Bonitomes/tree/main/scripts) to count the occurrence of alternative topologies centered around the *Mycoavidus* BRE clade A-CaGg node that our novel topology hinges on. 16S rDNA maximum likelihood phylogenies were created using the alignment, and trimming approaches described above. All trees were visualized with ggtree v3.6.2 in R (Yu et al. 2018; Yu et al. 2017). To evaluate statistical support for co-evolution, we analyzed phylogenies of fungal hosts and their endosymbionts using ParaFit implemented in R (Legendre et al., 2002).

2.5. Functional genome annotation and comparative genomics

The putative functions of predicted proteomes from free-living and endohyphal Burkholderiales (Table S1) were assigned using a variety of comparative genomic tools including PFAM and Gene Ontology (GO terms) were annotated with interproscan v5.56–89.0 using default options. To assess gene loss and the potential for newly sequenced *Mycoavidus* isolates to carry out core metabolic functions, we annotated and compared their predicted proteomes to the KEGG PATHWAY database using *kofamKOALA* (Aramaki et al., 2020). We used the resulting proteome annotations to assess completeness of amino acid and fatty acid metabolic pathways by leveraging pathway linearity and completeness information encoded in KEGG pathway module strings (Table S2) implemented with custom scripts available at https://github.com/amsesk/EHB-Bonitomes.git. These analyses resulted in presence-absence matrices indicating completeness by pathway step for serine, threonine, methionine, and histidine biosynthesis as well as beta-oxidation and pathways across our 70-taxon set. We visualized these matrices using *tidyverse* (Wickham et al., 2019) and *ggtree* in R.

To evaluate the presence of predicted secondary metabolite clusters in BRE genomes including *Mycoavidus* and free-living Burkholderiales, genomes were annotated with antiSMASH v6.0.1 (Blin et al., 2021), running searches against known clusters and subclusters, as well as the MiBIG database. We used custom scripts to parse outputs which are available at https://github.com/amsesk/antiSMASH_tools, and visualized cluster counts and their phylogenetic distributions using *ggplot2* and *ggtree* in R.

3. Results

3.1. Mycoavidus species genome assembly size and content

We assembled 14 draft genomes for *Mycoavidus* spp. associated with Mortierellomycotina fungi, that ranged from 2.1 to 3.2 Mbp in cumulative size (Table 1). Despite variable sequencing depth between metagenomic assemblies (mean coverage ranging from 2.87 to 130.80 x), our filtering approach yielded genomic assemblies that were composed of 1–235 contigs per bacterial isolate and estimated to be 82.99–92.8 7% complete according to BUSCO metrics based on the burkholderiales_odb10 database of core orthologous genes (Table 1) (Manni et al. 2021).

3.2. rDNA phylogenies support three major clades within monophyletic Mycoavidus

To resolve the phylogenetic placement of our newly sequenced *Mycoavidus* genomes, we inferred single locus phylogenetic trees based on 16S rDNA sequences extracted from genomic contigs. This phylogeny included assembled 16S rDNA loci from our 14 *Mycoavidus* draft genomes combined with an extensive sampling of publicly available 16S rDNA sequences for additional BRE and other free living Burkholderiales bacteria (Table 2). Our 16S rDNA phylogeny supports the three clades of *Mycoavidus* with 71%, 99%, and 73% bootstrap support for *Mycoavidus* BRE clades A, B, and C, respectively (Fig. 1). These results are consistent with previous reports that refer to these groups as sub-clades MorBRE A, B, and C (Okrasińska et al. 2021) or *M. cysteinexigens* clades A, B, and C

(Takashima et al. 2018). We refer to these groups here as *Mycoavidus* BRE clades A, B, and C. Our 16S rDNA phylogeny resolves an independent *Mycoavidus* BRE clade A (Fig. 1), and a sister clade of nematode associated BRE intracellular associates in the genus *Xiphinema* with 99% bootstrap support.

3.3. Genome-scale phylogenetic analyses suggest that Mycoavidus is a composed of several lineages

To investigate the evolutionary relationships between different *Mycoavidus* BRE clades using genome scale data, we constructed phylogenies based on 652 core orthologs present at sufficient occupancy across a 70-taxon set that sampled all major clades of free-living and endosymbiotic Burkholderiales and covered 187,114 amino acid positions (Table 1, Table S1). In the best ML tree, we resolve independent *Mycoavidus* BRE clades A, B, and C, which are well supported by our 16S rDNA phylogenies (Fig. 1). We recovered an overall topology demonstrating sisterhood between the clades *Burkholderia* - *Paraburkholderia* - *Mycetohabitans* and *Mycoavidus* - *Candidatus* Glomeribacter gigasporarum supported in other genome-scale phylogenies (Guo et al. 2020) and also recapitulated here (Fig. 2). However, in disagreement with past BRE phylogenies, our genome-scale phylogeny garners strong (100%) bootstrap support for a *Mycoavidus* related lineage composed of two independent clades separated by a medial-diverging CaGg (Fig. 2).

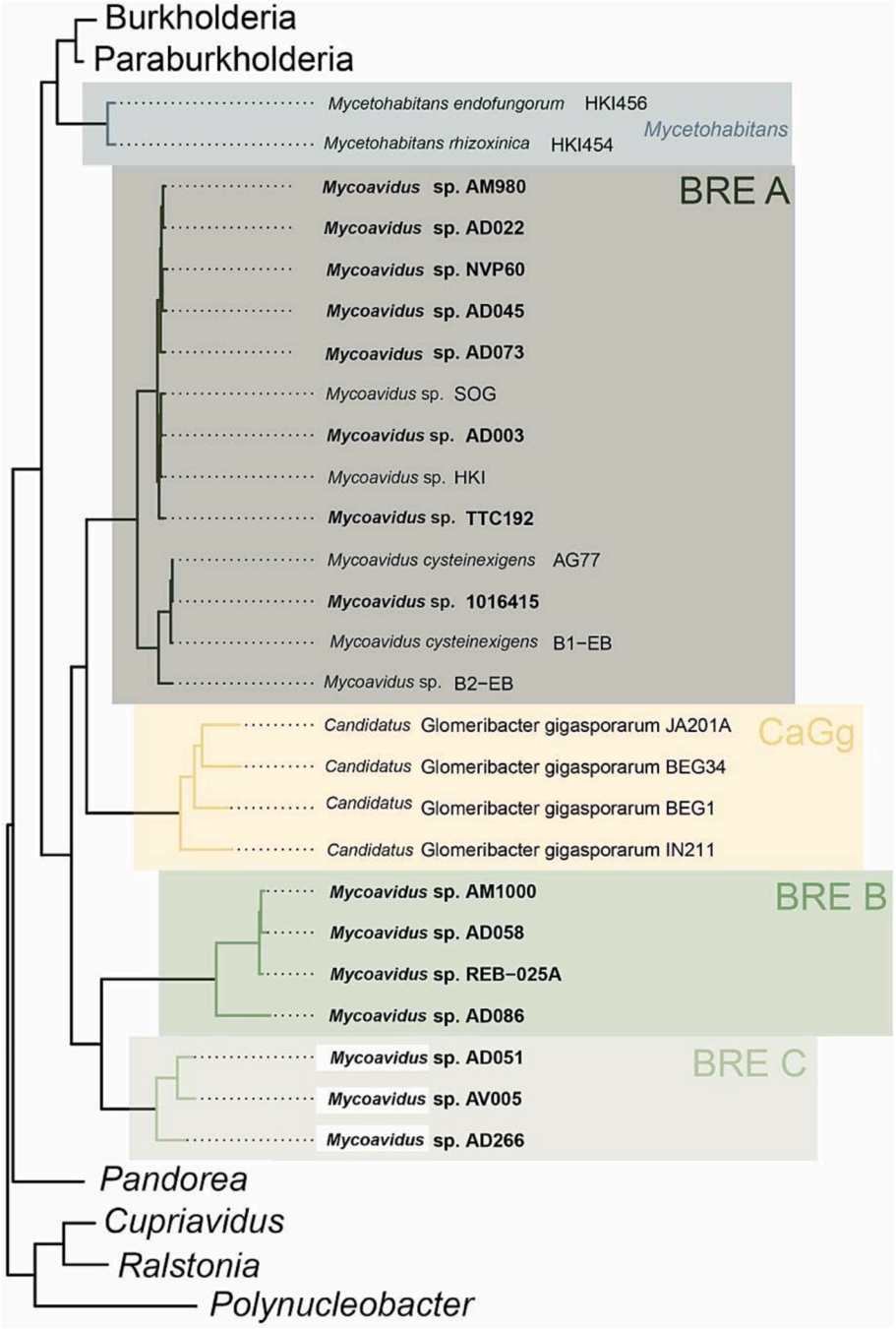

Fig. 2. Genome-scale phylogenomic tree of Burkholderiaceae including novel *Mycoavidus* isolates. This tree is based on 652 core orthologous genes from the *burkholderiales_odb10* database. Concatenated alignment covers 187,144 amino acid positions with marker occupancy of 94.23% across 70 bacterial genomes from Burkholderiaceae. Major groups of *Mycoavidus* BRE are highlighted by colored boxes and labeled. New genomes generated herein are bolded.

We resolve *Mycoavidus* BRE clade A and CaGg as sister lineages, with *Mycoavidus* BRE clades B and C sister to clade A-CaGg with 100% bootstrap support.

To further investigate support for the relationship of CaGg and *Mycoavidus* BRE clade A, and because the utility of bootstrapping can decline with increasing alignment size, we calculated gene tree concordance factors (gCF) at important nodes. We also computed an ASTRAL consensus tree based on all 652 gene trees. The ASTRAL tree resolved the same topology in *Mycoavidus* BRE clade A-CaGg as the best ML tree (Figure S1). Calculated support values for the operative nodes were 46% and 53% for gCF (Figure S2) and quartet support (ASTRAL-III), respectively (Figure S3). We used custom scripts to identify and count the frequency of competing topologies at this node within our set of 652 gene trees. We also conducted an approximately unbiased (AU) test for alternative placement of CaGg that would result in a monophyletic *Mycoavidus* in IQTree2, but found that it was significantly worse than the best tree (pAU = 1.14e-39). We algorithmically identified the topological relatedness of tips belonging to *Mycoavidus* BRE clades A, B, C, and CaGg in each gene tree and counted their frequency across the set of 652 gene trees. In concordance with gCF and ASTRAL results, the topology represented in the best ML tree is the most frequent among complete 4-tip topologies represented in the gene trees (203/507 resolved 4-tip topologies) (Figure S4). The next most frequent topology, which constitutes 63/507 of resolved 4-tip topologies, also disagrees with the currently accepted topology resolved by 16S rDNA trees and resolves *Mycoavidus* BRE clades B/C and CaGg as sister lineages, with *Mycoavidus* BRE clade A as sister to that group (Figure S4). The currently accepted topology, based on 16S rDNA trees (i.e., monophyletic *Mycoavidus* BRE clades A and B with C sister to CaGg), is supported in only 51/507 resolved 4-tip topologies. Other alternative 4-tip topologies occur at lower frequencies ranging from 5 to 34 of the 507 resolved topologies (Figure S4). The sister relationship between *Mycoavidus* BRE clade A and CaGg is resolved in topological contexts outside that of the best ML topology, totaling 286 of the 636 (~45%) gene tree topologies resolved to >=2 tip occupancy.

3.4. Fungal host association specificity varies by Mycoavidus clade

To evaluate the host-specificity of *Mycoavidus* spp. by lineage we created a host fungal phylogeny and compared associations of new and existing *Mycoavidus* isolates by host associations in each clade. We observed that the majority of the 13 isolates in *Mycoavidus* BRE clade A are associated with *Linnemannia* spp., particularly with *L. elongata* and *L. gamsii* (Fig. 3, Table 1). Five isolates were outliers to this pattern, *Mycoavidus* sp. B2-EB, *Mycoavidus* AD003, *Mycoavidus* HKI, *Mycoavidus* TTC192 and *Mycoavidus* SOG that were associated with *Entomortierella, Podila humilis, P. verticillata,* and *Stylopage* (Zoopagomycota) respectively (Table 1, Fig. 3). The most divergent fungal-endosymbiont pair was *Mycoavidus* SOG which was detected in *Sytlopage* (Zoopagomycota) which is a new host for *Mycoavidus*. *Mycoavidus* BRE clade B isolates exhibited the highest degree of specificity, associating only with *Podila* spp. (Fig. 3, Table 1). In contrast, *Mycoavidus* BRE clade C isolates

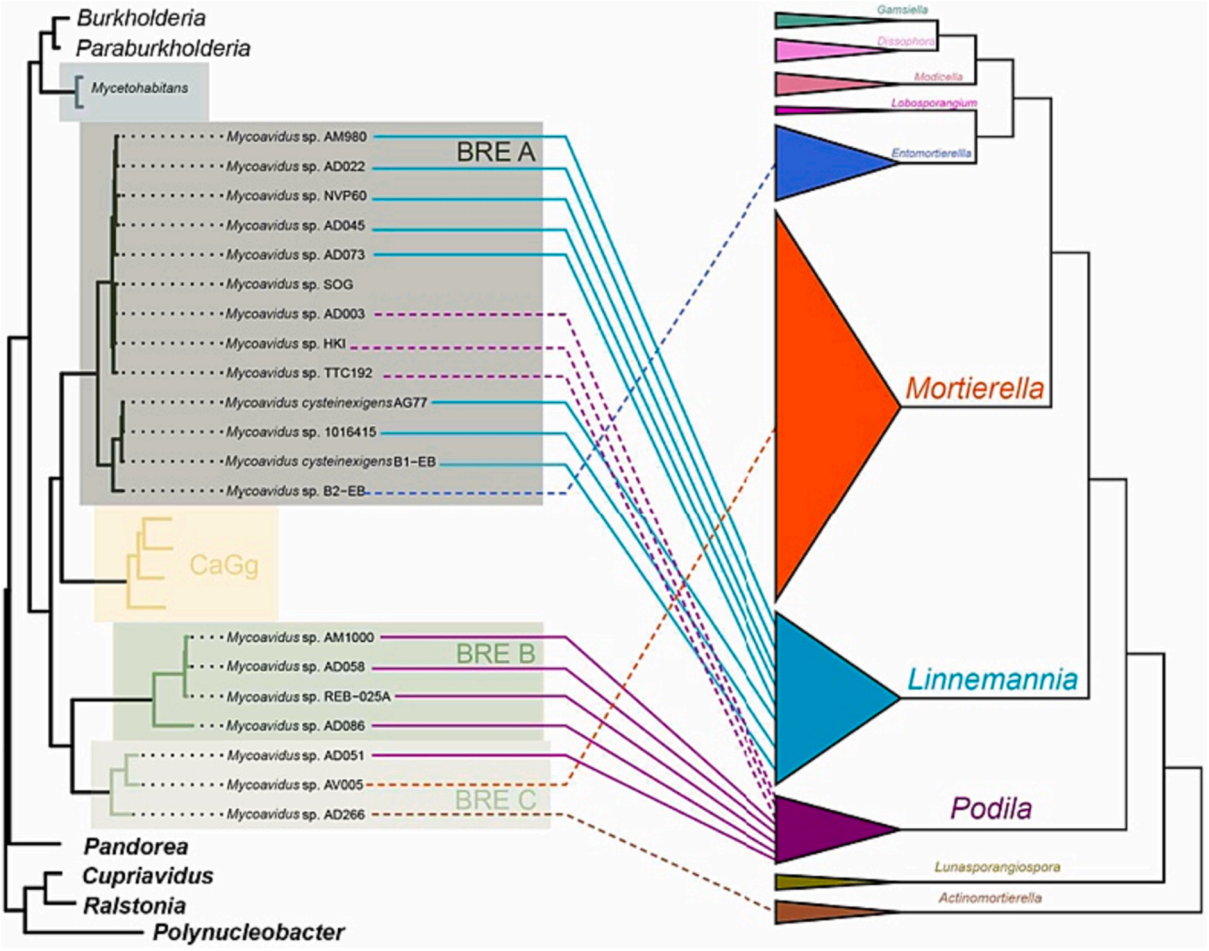

Fig. 3. Best concatenated genome-scale phylogenetic tree for Burkholderiaceae (i.e., Fig. 2) opposite a cladogram containing several important host genera in the Mortierellomycotina, adapted from Vandepol et al. 2020. Intersecting lines between the endosymbiont (left) or host (right) trees indicate known associations. The continuity of intersecting lines indicates if each known association significantly differs from the expectation of independent evolution and is cophylogenetic (solid) or not (dashed).

exhibited the least specificity, associating with unique hosts. Within this group, *Mycoavidus* AD051, AV005, and AD266 were found with *Podila, Actinomortierella,* and *Mortierella* species, respectively (Fig. 3, S5 Table 1).

4. *Mycoavidus* genomes contain unique biosynthetic gene clusters relative to their free-living ancestors

To evaluate potential host beneficial functionality encoded in *Mycoavidus* genomes we annotated putative biosynthetic gene clusters (BGCs) in our newly sequenced genomes. Compared to their free-living relatives, we found that *Mycoavidus* BRE clades contain genes predicted to encode unique types of nonribosomal peptide synthases (NRPSs) and *trans*-AT polyketide synthases (transAT-PKSs). We identified two classes of cytotoxic NRPSs (rhizomides and luminmides) potentially encoded by *Mycoavidus* isolates that had no analogs in the free-living Burkholderiales that we analyzed (Figure S6). A large portion of these NRPSs (73%) shared high cluster similarity (100%) with known rhizomides, a class of short-chain cyclic peptides. In order to assess the distribution of the putative rhizomide-like BGCs across the *Burkholderia*-related endosymbionts sequenced in this study, we mapped their predicted genetic structure (as PFAM domains) onto our genome-scale phylogenetic tree (Fig. 4). We found that rhizomide-like gene clusters, similar to those previously detected by BGC mining in *M. rhizoxinica* HKI454 (Ouyang et al., 2020), are distributed across *Mycoavidus* clades and are present in 46%, 25%, and 67% of *Mycoavidus* BRE clades A, B, and C isolates respectively (Fig. 4). Knowing the effect that assembly fragmentation can have on annotation accuracy, we included both the fragmented and entire genomes for *Mycoavidus* isolate AG77 and excluded rhizomide-like NRPSs predicted within < 500 bp of the ends of contigs.

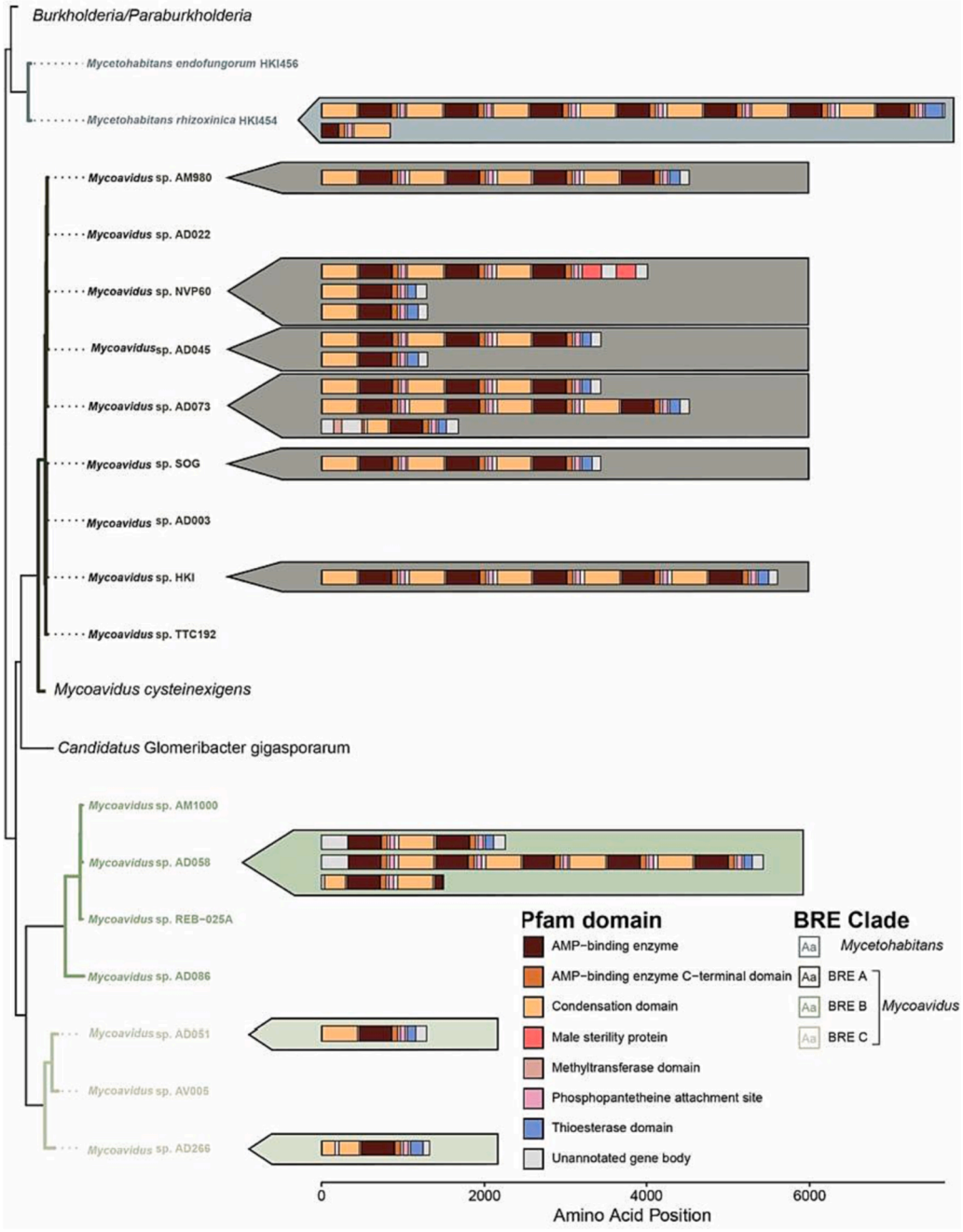

Fig. 4. Modified genome-scale phylogenetic tree of Burkholderiaceae isolates including novel *Mycoavidus* BRE isolates showing the distribution of predicted rhizomide-like nonribosomal peptide synthases in BRE genomes. The repetitive architecture of the locus is shown as colored boxes indicating the best interproscan hit of regions to functional domains contained in the PFAM database. *Mycoavidus* BRE clade A has been collapsed due to the absence of annotated rhizomides in select isolates.

5. Endosymbiosis shapes *Mycoavidus* genome size and content

To evaluate the degree of genome reduction in *Mycoavidus* and predict the effect that gene loss could have on metabolic function, we quantified the presence of genes required for amino acid and lipid metabolism, as defined in KEGG PATHWAYS (Table S2). We identified common and unique loss patterns between and within *Mycoavidus* clades (Fig. 5). We found that all newly sequenced and assembled *Mycoavidus* isolates share reduction in some pathways, while other patterns of gene loss and pathway inactivation patterns are lineage specific (Fig. 5). For example, *Mycoavidus* BRE clades A and C share losses in histidine biosynthesis while *Mycoavidus* BRE clade B is differentiated by losses in the threonine pathway (Fig. 5). Meanwhile, all sequenced *Mycoavidus* BRE clades and CaGg clades show degraded pathways for cysteine biosynthesis (Fig. 5). We also noted the absence of genes involved in the catabolism of fatty acids their distribution was asymmetrical across *Mycoavidus,* and the loss or retention of orthologues in the beta-oxidation pathway differs consistently between *Mycoavidus* BRE clades and *Candidatus* Glomeribacter gigasporarum (Fig. 5).

6. Discussion

6.1. Novel Mycoavidus genomes illuminate biology and evolutionary history of fungal endosymbionts

After sorting and assembling endosymbiont genomes from mixed fungal and endosymbiont raw genome sequencing data, we obtained 14 *Mycoavidus* genomes that we estimated to be 82.99–92.87% complete compared to BUSCO orthologues (Table 1). A phylogenetic analysis of 16S rDNA sequences extracted from those genomes and existing isolates demonstrate their relationship to one another and improve our understanding of the phylogenomic diversity within the group. We observed 71%, 99%, and 73% bootstrap support for the three distinct clades of *Mycoavidus* (Fig. 1). The clades mirror topology of *Mycoavidus* clades A, B, and C in previously reported work (Takashima et al. 2018; Okrasińska et al. 2021). These results indicate there is extensive molecular diversity within *Mycoavidus* that may warrant the description of new taxa. In parallel, we analyzed genome scale phylogenetic data of the newly sequenced *Mycoavidus* isolates presented here (Table 1) and existing BRE genomes of *Candidatus* Glomeribacter, *Mycetohabitans,* and free-

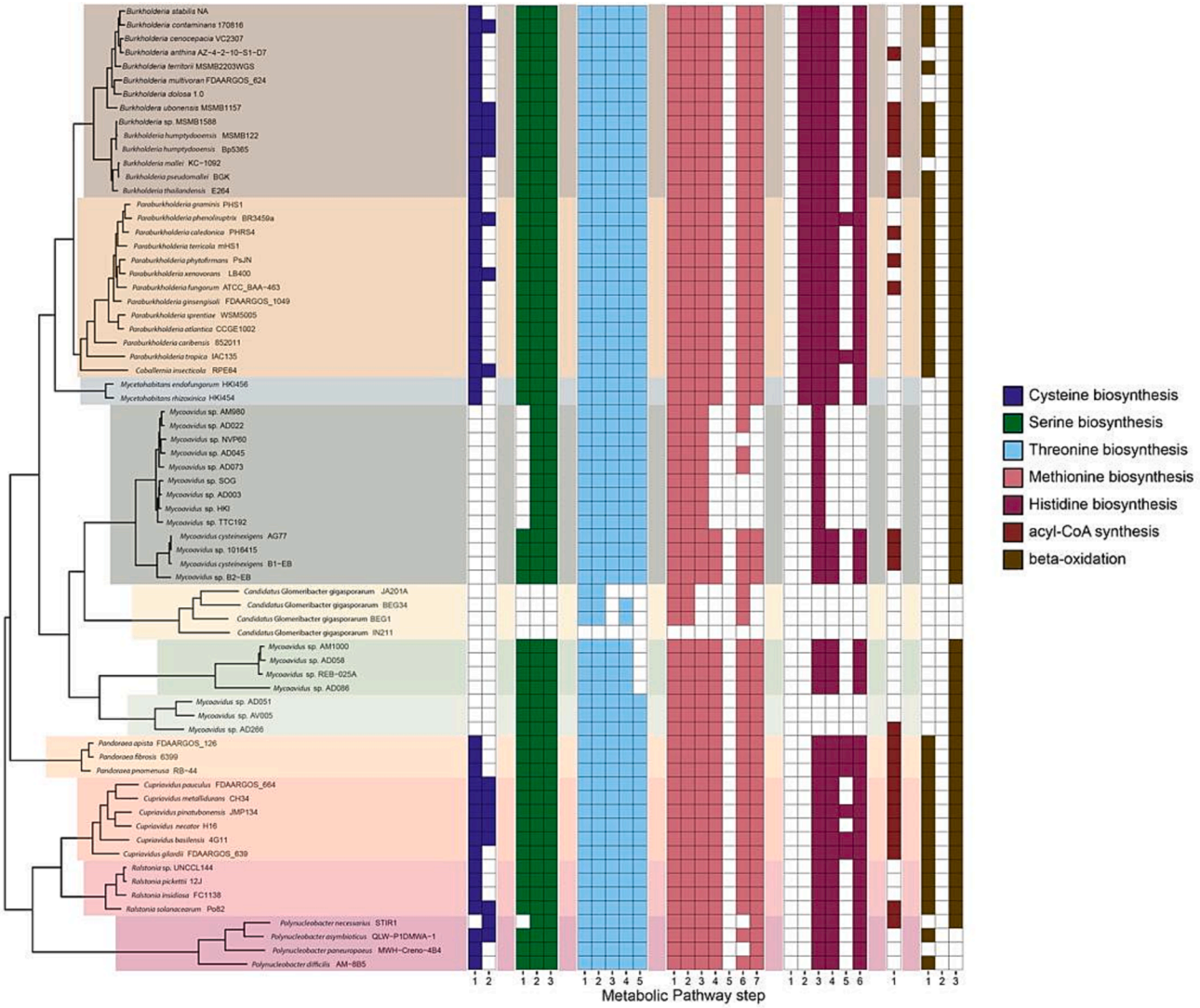

Fig. 5. Phylogenetically clustered binary heatmaps showing the presence of genes involved in amino acid biosynthesis or beta-oxidation as columns in the predicted proteomes of BRE including novel Mycoavidus isolates as rows. Colored boxes indicate the presence of proteins within bacterial genomes to encode the requisite enzymes for each step, while white boxes include the pathway step cannot be completed based on the proteome. Presence of proteins in pathways were assessed by the homology of predicted proteins to associated KEGG Ontology terms associated with each pathway in the KEGG MODULE database (see Table S2). Based on KEGG module strings, each step in the pathways is sequential toward completion of the synthesis or catabolism pathway.

living *Burkholderia* relatives (Table S1). While we still observed internal support for clades *Mycoavidus* BRE clades A-C, the phylogenomic tree topology portrays *Mycoavidus* BRE clades B and C sister to a monophyletic lineage comprising CaGg and *Mycoavidus* BRE clade A, which is sister to BRE A-CaGg (Fig. 2). This finding implies that *Mycoavidus* requires taxonomic revision. Conversely, once more taxa related to *Mycoavidus* are described synonymizing taxa may also be warranted in accordance with the International Code of Botanical Nomenclature. In addition, the stablilizing taxonomic landscape on the fungal host side will contribute to our understanding of specificity in these symbioses. In contrast, our 16S rDNA phylogeny (Fig. 1), and previously published work based on 16S rDNA, portray CaGg as a well-supported sister group to a monophyletic *Mycoavidus* (Sharmin et al., 2018; Uehling et al., 2017). Previous studies have documented substantial molecular diversity among 16S sequences of CaGg (Desirò et al., 2014), but so far none have suggested that *Mycoavidus* and related lineages may be ancestral to CaGg.

After analyzing endosymbiont host associations within currently available *Mycoavidus* BRE clades, we found that clades vary in host specificity (Fig. 3). *Mycoavidus* BRE clades A and B isolates are more host-specific, associated primarily with *Linnemannia* and *Podila* spp. In contrast, *Mycoavidus* BRE clade C isolates were obtained from diverse host genera (Fig. 3). Considered together with the lineage-specific divergent evolutionary history and functional genomic capacity demonstrated here, these data indicate endosymbiont biology, host specificity, and transmission vary by endosymbiont clade. However, these results should be contextualized in that *Mycoavidus* BRE clades B and C are currently under sampled (containing 4 and 3 bacterial isolates respectively) compared to clade A which contains 13 isolates, and that Mortierellomycotina fungi continue to undergo rapid taxonomic reorganizations (Vandepol et al., 2020) that influence the specificity analyses. In addition, we show *Mycoavidus* associates with Zoopagomycota fungal hosts, indicating the host range for these bacterial fungal interactions may be broader that initially understood. The future sequencing of more *Mycoavidus* genomes, generation of functional interaction data to test these hypotheses and the evaluation of transmission mechanisms will elucidate key pieces of microbial biology in these symbiotic systems.

6.2. *Mycoavidus encode predicted rhizomide-like secondary metabolites*

To evaluate the potential secondary metabolic potential in endosymbiotic and free-living Burkholderiaceae genomes (Table 1, Table S1), we predicted BGCs in 26 fungal endosymbionts including our newly sequenced *Mycoavidus* isolates and 44 Burkholderiaceae genomes representative of free-living diversity in this group. We found that *Mycoavidus* and related endosymbionts contain unique putative NRPS and transAT-PKS classes compared to the free-living Burkholderiales we analyzed (Fig. 5, S6). Of particular interest were predicted NRPS BGCs with high sequence similarity to rhizomides which were present in some of the *Mycoavidus* genomes, though their presence and copy number varied by bacterial isolate and clade (Fig. 4). Rhizomides have demonstrated antitumor activity against human cell lines for gastric, breast, liver, cervical and lung cancers and antimicrobial activity against cucumber downy mildew, *Bacillus* and *Staphylococcus* bacteria (Wang et al., 2018). Although we did not isolate or evaluate function of rhizomides, luminmides or other secondary metabolites described here, it is well known that endosymbiont derived secondary metabolites can confer protection for fungal hosts and their symbiotic inhabitants against predators and soil dwelling competitors (Büttner et al., 2021; Flórez et al., 2017; Itabangi et al., 2022; Scherlach et al., 2006). Further, given that BGCs modify amino acids that are host derived, the presence of these gene clusters implies that fungal and endosymbiont primary and secondary metabolism are closely intertwined, underpinning these symbioses. We hypothesize these metabolites shape ecological interactions between fungi and their host organisms and their evolution. Future studies evaluating the abundance, diversity, and function of secondary metabolites in fungal-endosymbiont systems will shed light on habitat specific benefits endosymbionts give their fungal hosts.

6.3. *Endosymbiont genomes adapt to intracellular life*

We evaluated the relative genome size of endosymbionts compared to their free-living bacterial relatives and observed that all *Mycoavidus* genomes sequenced to date are extraordinarily streamlined, being half to a third the size of their free-living relatives (Tables 1, Table S1). We also investigated ubiquity of gene loss and retention in *Mycoavidus* amino acid biosynthesis and fatty acid metabolic pathways and documented both ubiquitous and lineage specific patterns (Fig. 5). Regarding amino acid biosynthesis, we note that while all *Mycoavidus* clades and CaGg are missing entire pathways such as cysteine biosynthesis. There are also lineage-specific patterns of loss and retention in threonine, methionine, and histidine biosynthetic pathways (Fig. 5). Given the unique molecular identity (Figs. 1, 2) and distinct geographic origin of these isolates (Table 1), this pattern demonstrates convergent genome contraction and gene loss in common functional pathways. Our results are potentially limited by the varying degrees of genome fragmentation in our assemblies. However, we assessed how fragmentation impacts functional annotation using the fragmented and entire assemblies for *Mycoavidus* AG77 and observed striking consistency in our phylogenomic, secondary metabolic, and metabolic pathway analyses (Figs. 2, 3, 5) between fragmented and entire genomes in these analyses. Coupled with their high BUSCO completeness estimates (Table 1), we feel confident these data accurately reflect evolutionary patterns.

With regards to fatty acid metabolism, we observed that functional aspects of fatty acid catabolic potential are conserved in select bacterial isolates, and that BRE generally have limited beta-oxidation in comparison with their free-living relatives (Fig. 5). We surveyed for the presence of orthologues to acyl-CoA synthesis (Table S2), which produce a group of co-enzymes that metabolize fatty acids (Weete, 2012). We also searched the new *Mycoavidus* genomes for orthologs of genes in the beta-oxidation pathway (Table S2), which is a process in which fatty acids are catabolized to generate acetyl-CoA that can enter the citric acid cycle (Weete, 2012). We found that in all available *Mycoavidus* isolates, the third gene in the pathway (Table S2) is conserved while the first two are missing. We also note this is the only consistent loss and retention pattern differentiating *Candidatus* Glomeribacter gigasporarum and *Mycoavidus* (Fig. 5). As previously reported, select bacterial isolates in *Mycoavidus* BRE clade A including isolate AG77 are equipped to catabolize fatty acids (Li et al., 2017; Uehling et al., 2017). However, these results suggest that *Mycoavidus* lineages and isolates differ in the extent of their ability to utilize fatty acids to generate acetyl-CoA. Future investigations elucidating shared and unique molecular mechanisms of endosymbiont host interactions by fungal lineage or isolate will shed light on forces driving macro-evolutionary trends in this system.

The convergent evolution observed here may be driven by the commonality of host provisioned resources (Li et al., 2017; Uehling et al., 2017). It is well known that the environment shapes bacterial genome evolution (McCutcheon et al., 2019), with genome size and content reflecting how selective pressures differ by unique habitat. Host-derived selective pressures such as provisioning of amino acids have the potential to shape the patterns we observe here. By noting convergent gene loss in functionally similar pathways that are commonly degraded in endosymbionts from closely related hosts, we gain direct insight into shared resources between hosts and endosymbionts. This pattern holds across eukaryotic endosymbionts including insects and intracellular parasites (Wernegreen, 2015). Fungal and insect endosymbionts differ in several key aspects that likely shape their evolutionary trajectories distinctly. Both fungi and arthropods contain bacteria that alter host reproductive biology (Harcombe and Hoffmann, 2004; Mondo et al., 2017; Takashima et al., 2020) and many endosymbionts are resistant to aseptic culture techniques. Yet, fungal hosts are regularly found without

bacteria (Okrasińska et al., 2021; Sharmin et al., 2018), can be cleared of endosymbionts with antibiotics, and some appear to be horizontally transferred between diverse hosts. Even considering these notable differences in eukaryotic endosymbiont systems, the shared characteristics of genome contraction and functional convergence through gene loss suggest that the eukaryotic intracellular environment shapes endosymbionts genomes in patterned and predictable rules of life.

Funding information

This work was supported by NSF 2202410 and NSF 2030338 to JU, TP, RA, EB and JK, DOE SFA LANLF59T and NSF 1737898 to GB, BBSRC BB/W002760/1 to EB and JK, and NRF 132803 to RA. The work (proposals: https://doi.org/10.46936/10.25585/60001028 and https://doi.org/10.46936/10.25585/60001291) conducted by the U.S. Department of Energy Joint Genome Institute (https://ror.org/04xm1d337), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231.

CRediT authorship contribution statement

Kevin Amses: Data curation, Methodology, Formal analysis, Data visualization, Original draft writing. **Alessandro Desiró:** Strain isolation, DNA extraction, Review, and editing. **Abigail Bryson:** Strain isolation, DNA extraction, Review, and editing. **Igor Grigoriev:** Project administrator bioinformatics, Review, and editing. **Stephen Mondo:** Data curation, Review, and editing. **Anna Lipzen:** Data curation, Formal analysis, Review, and editing. **Kurt LaButti:** Data curation, Formal analysis, Review, and editing. **Robert Riley:** Data curation, Formal analysis. **Vasanth Singan:** Data curation, Formal analysis. **Paris Salazar-Hamm:** Data curation, Formal analysis. **Jason King:** Funding acquisition, Review, and editing. **Elizabeth Ballou:** Funding acquisition, Review, and editing. **Teresa Pawlowska:** Funding acquisition, Review, and editing. **Rasheed Adeleke:** Funding acquisition, Review, and editing. **Gregory Bonito:** Project administration, Funding acquisition, Conceptualizing, Supervision, Writing, and editing. **Jessie Uehling:** Conceptualization, Methodology, Project administration, Funding acquisition, Methodology, Supervision, Original draft writing, Review, and editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank Drs. Andy Miller, Kerry O'Donnel and Porras Alfaro for contributing fungal strains containing endosymbionts, and Megan Korne for assistance in submitting isolates to the Westerdijk Fungal Biodiversity Institute (CBS) fungal culture collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fgb.2023.103838.

References

Afri, H., Yong, G., Yusuke, T., Kazuhiko, N., Hiroyuki, O., Tomoyasu, N., 2022. Whole-Genome Sequence of Entomortierella parvispora E1425, a Mucoromycotan Fungus Associated with Burkholderiaceae-Related Endosymbiotic Bacteria. Microbiology Resource Announcements 11, e01101–e2021.

Amses, K.R., Davis, W.J., James, T.Y., 2020. SCGid, a consensus approach to contig filtering and genome prediction from single cell sequencing libraries of uncultured eukaryotes. Bioinformatics 36, 1194–2000.

Aramaki, T., Blanc-Mathieu, R., Endo, H., Ohkubo, K., Kanehisa, M., Goto, S., Ogata, H., 2020. KofamKOALA: KEGG Ortholog assignment based on profile HMM and adaptive score threshold. Bioinformatics 36, 2251–2252.

Bennett, G.M., Moran, N.A., 2013. Small, smaller, smallest: the origins and evolution of ancient dual symbioses in a Phloem-feeding insect. Genome Biol. Evol. 5, 1675–1688.

Bianciotto, V., Lumini, E., Bonfante, P., Vandamme, P., 2003. "Candidatus Glomeribacter gigasporarum"gen. nov., sp. nov., an endosymbiont of arbuscular mycorrhizal fungi. Int. J. Syst. Evol. Microbiol. 53, 121–124.

Blin, K., Shaw, S., Kloosterman, A.M., Charlop-Powers, Z., van Wezel, G.P., Medema, M. H., Weber, T., 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res. 49, W29–W35.

Bonfante, P., Desirò, A., 2017. Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. ISME J. 11, 1727–1735.

Büttner, H., Niehs, S.P., Vandelannoote, K., Cseresnyés, Z., Dose, B., Richter, I., Gerst, R., Figge, M.T., Stinear, T.P., Pidot, S.J., Hertweck, C., 2021. Bacterial endosymbionts protect beneficial soil fungus from nematode attack. PNAS 118. https://doi.org/10.1073/pnas.2110669118.

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

Chong, R.A., Park, H., Moran, N.A., 2019. Genome Evolution of the Obligate Endosymbiont Buchnera aphidicola. Mol. Biol. Evol. 36, 1481–1489.

Denisov, G., Walenz, B., Halpern, A.L., Miller, J., Axelrod, N., Levy, S., Sutton, G., 2008. Consensus generation and variant detection by Celera Assembler. Bioinformatics 24, 1035–1040.

Desirò, A., Salvioli, A., Ngonkeu, E.L., Mondo, S.J., Epis, S., Faccio, A., Kaech, A., Pawlowska, T.E., Bonfante, P., 2014. Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi. ISME J. 8, 257–270.

Desirò, A., Takashima, Y., Bonito, G., Nishizawa, T., Narisawa, K., Bonfante, P., 2023. Investigating Endobacteria that Thrive Within Mucoromycota. Methods Mol. Biol. 2605, 293–323.

Flórez, L.V., Scherlach, K., Gaube, P., Ross, C., Sitte, E., Hermes, C., Rodrigues, A., Hertweck, C., Kaltenpoth, M., 2017. Antibiotic-producing symbionts dynamically transition between plant pathogenicity and insect-defensive mutualism. Nat. Commun. 8, 15172.

Freire, B., Ladra, S., Parama, J.R., 2022. Memory-Efficient Assembly Using Flye. IEEE/ACM Trans. Comput. Biol. Bioinf. 19, 3564–3577.

Gryganskyi, A.P., Golan, J., Dolatabadi, S., Mondo, S., Robb, S., Idnurm, A., Muszewska, A., Steczkiewicz, K., Masonjones, S., Liao, H.-L., Gajdeczka, M.T., Anike, F., Vuek, A., Anishchenko, I.M., Voigt, K., de Hoog, G.S., Smith, M.E., Heitman, J., Vilgalys, R., Stajich, J.E., 2018. Phylogenetic and Phylogenomic Definition of Rhizopus Species. G3: Genes, Genomes, Genetics. G3 (8), 2007–2018.

Guo, Y., Takashima, Y., Sato, Y., Narisawa, K., Ohta, H., Nishizawa, T., 2020. Mycoavidus sp. strain B2-EB: comparative genomics reveals minimal genomic features required by a cultivable Burkholderiaceae-related endofungal bacterium. Appl. Environ. Microbiol 86 (18), e01018–20.

Harcombe, W., Hoffmann, A.A., 2004. Wolbachia effects in Drosophila melanogaster: in search of fitness benefits. J. Invertebr. Pathol. 87, 45–50.

Ingrid, R., Silvia, R., Zoltán, C., Iuliia, F., Hannah, B., Niehs, S.P., Ruman, G., Kirstin, S., Thilo, F.M., Falk, H., Christian, H., 2022. Toxin-Producing Endosymbionts Shield Pathogenic Fungus against Micropredators. MBio 13, e01440–e10522.

Itabangi, H., Sephton-Clark, P.C.S., Tamayo, D.P., Zhou, X., Starling, G.P., Mahamoud, Z., Insua, I., Probert, M., Correia, J., Moynihan, P.J., Gebremariam, T., Gu, Y., Ibrahim, A.S., Brown, G.D., King, J.S., Ballou, E.R., Voelz, K., 2022. A bacterial endosymbiont of the fungus Rhizopus microsporus drives phagocyte evasion and opportunistic virulence. Curr. Biol. 32, 1115–1130.e6.

Kang, D.D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., Wang, Z., 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7, e7359.

Kronenberg, Z.N., Hall, R.J., Hiendleder, S., Smith, T.P.L., Sullivan, S.T., Williams, J.L., Kingan, S.B., 2018. FALCON-Phase: integrating PacBio and Hi-C data for phased diploid genomes. BioRxiv, 327064.

Lackner, G., Moebius, N., Partida-Martinez, L., Hertweck, C., 2011. Complete genome sequence of Burkholderia rhizoxinica, an endosymbiont of Rhizopus microsporus. J. Bacteriol. 193, 783–784.

Legendre, P., Desdevises, Y., Bazin, E., 2002. A statistical test for host-parasite coevolution. Syst. Biol. 51, 217–234.

Li, Z., Yao, Q., Dearth, S.P., Entler, M.R., Castro Gonzalez, H.F., Uehling, J.K., Vilgalys, R. J., Hurst, G.B., Campagna, S.R., Labbé, J.L., Pan, C., 2017. Integrated proteomics and metabolomics suggests symbiotic metabolism and multimodal regulation in a fungal-endobacterial system. Environ. Microbiol. 19, 1041–1053.

Liao, H.-L., Bonito, G., Rojas, J.A., Hameed, K., Wu, S., Schadt, C.W., Labbé, J., Tuskan, G.A., Martin, F., Grigoriev, I.V., Vilgalys, R., 2019. Fungal Endophytes of Populus trichocarpa Alter Host Phenotype, Gene Expression, and Rhizobiome Composition. Mol. Plant Microbe Interact. 32, 853–864.

Luginbuehl, L.H., Menard, G.N., Kurup, S., Van Erp, H., Radhakrishnan, G.V., Breakspear, A., Oldroyd, G.E.D., Eastmond, P.J., 2017. Fatty acids in arbuscular mycorrhizal fungi are synthesized by the host plant. Science 356, 1175–1178.

McCutcheon, J.P., Boyd, B.M., Dale, C., 2019. The life of an insect Endosymbiont from the cradle to the grave. Curr. Biol. 29, R485–R495.

Misztal, P.K., Lymperopoulou, D.S., Adams, R.I., Scott, R.A., Lindow, S.E., Bruns, T., Taylor, J.W., Uehling, J., Bonito, G., Vilgalys, R., Goldstein, A.H., 2018. Emission Factors of Microbial Volatile Organic Compounds from Environmental Bacteria and Fungi. Environ. Sci. Tech. 52, 8272–8282.

Mondo, S.J., Toomer, K.H., Morton, J.B., Lekberg, Y., Pawlowska, T.E., 2012. Evolutionary stability in a 400-million-year-old heritable facultative mutualism. Evolution 66, 2564–2576.

Mondo, S.J., Lastovetsky, O.A., Gaspar, M.L., Schwardt, N.H., Barber, C.C., Riley, R., Sun, H., Grigoriev, I.V., Pawlowska, T.E., 2017. Bacterial endosymbionts influence host sexuality and reveal reproductive genes of early divergent fungi. Nat. Commun. 8, 1843.

Moran, N.A., 1996. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. PNAS 93, 2873–2878.

Ohshima, S., Sato, Y., Fujimura, R., Takashima, Y., Hamada, M., Nishizawa, T., Narisawa, K., Ohta, H., 2016. Mycoavidus cysteinexigens gen. nov., sp. nov., an endohyphal bacterium isolated from a soil isolate of the fungus Mortierella elongata. Int. J. Syst. Evol. Microbiol. 66, 2052–2057.

Okrasińska, A., Bokus, A., Duk, K., Gęsiorska, A., Sokołowska, B., Miłobędzka, A., Wrzosek, M., Pawłowska, J., 2021. New Endohyphal Relationships between Mucoromycota and Burkholderiaceae Representatives. Appl. Environ. Microbiol. 87 https://doi.org/10.1128/AEM.02707-20.

Ouyang, Q., Wang, X., Zhang, N., Zhong, L., Liu, J., Ding, X., Zhang, Y., Bian, X., 2020. Promoter Screening Facilitates Heterologous Production of Complex Secondary Metabolites in Burkholderiales Strains. ACS Synth. Biol. 9, 457–460.

Partida-Martinez, L.P., Groth, I., Schmitt, I., Richter, W., Roth, M., Hertweck, C., 2007. Burkholderia rhizoxinica sp. nov. and Burkholderia endofungorum sp. nov., bacterial endosymbionts of the plant-pathogenic fungus Rhizopus microsporus. International Journal of Systematic and Evolutionary Microbiology 57, 2583–2590.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490.

Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., Korobeynikov, A., 2020. Using SPAdes De Novo Assembler. Curr. Protoc. Bioinformatics 70, e102.

Sato, Y., Narisawa, K., Tsuruta, K., Umezu, M., Nishizawa, T., Tanaka, K., Yamaguchi, K., Komatsuzaki, M., Ohta, H., 2010. Detection of betaproteobacteria inside the mycelium of the fungus Mortierella elongata. Microbes Environ. 25, 321–324.

Scherlach, K., Partida-Martinez, L.P., Dahse, H.-M., Hertweck, C., 2006. Antimitotic Rhizoxin Derivatives from a Cultured Bacterial Endosymbiont of the Rice Pathogenic Fungus Rhizopus microsporus. J. Am. Chem. Soc. 128, 11529–11536.

Seemann, T., 2015. Barrnap. Github.

Sharmin, D., Guo, Y., Nishizawa, T., Ohshima, S., Sato, Y., Takashima, Y., Narisawa, K., Ohta, H., 2018. Comparative genomic insights into endofungal lifestyles of two bacterial endosymbionts, Mycoavidus cysteinexigens and *Burkholderia rhizoxinica*. Microbes Environ. 33 (1), 66–76. https://doi.org/10.1264/jsme2.me17138.

Spatafora, J.W., Chang, Y., Benny, G.L., Lazarus, K., Smith, M.E., Berbee, M.L., Bonito, G., Corradi, N., Grigoriev, I., Gryganskyi, A., James, T.Y., O'Donnell, K., Roberson, R.W., Taylor, T.N., Uehling, J., Vilgalys, R., White, M.M., Stajich, J.E., 2016. A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. Mycologia 108, 1028–1046.

Takashima, Y., Seto, K., Degawa, Y., Guo, Y., Nishizawa, T., Ohta, H., Narisawa, K., 2018. Prevalence and Intra-Family Phylogenetic Divergence of Burkholderiaceae-Related Endobacteria Associated with Species of Mortierella. Microbes Environ. 33, 417–427.

Takashima, Y., Degawa, Y., Nishizawa, T., Ohta, H., Narisawa, K., 2020. Aposymbiosis of a Burkholderiaceae-Related Endobacterium Impacts on Sexual Reproduction of Its Fungal Host. Microbes Environ. 35 https://doi.org/10.1264/jsme2.ME19167.

Telagathoti, A., Probst, M., Peintner, U., 2021. Habitat, Snow-Cover and Soil pH, Affect the Distribution and Diversity of Mortierellaceae Species and Their Associations to Bacteria. Frontiers in Microbiology 12, 669784.

Uehling, J., Gryganskyi, A., Hameed, K., Tschaplinski, T., Misztal, P.K., Wu, S., Desirò, A., Vande Pol, N., Du, Z., Zienkiewicz, A., Zienkiewicz, K., Morin, E., Tisserant, E., Splivallo, R., Hainaut, M., Henrissat, B., Ohm, R., Kuo, A., Yan, J., Lipzen, A., Nolan, M., LaButti, K., Barry, K., Goldstein, A.H., Labbé, J., Schadt, C., Tuskan, G., Grigoriev, I., Martin, F., Vilgalys, R., Bonito, G., 2017. Comparative genomics of Mortierella elongata and its bacterial endosymbiont Mycoavidus cysteinexigens. Environ. Microbiol. 19, 2964–2983.

Uehling, J.K., Salvioli, A., Amses, K.R., Partida-Martínez, L.P., Bonito, G., Bonfante, P., 2023. Bacterial endosymbionts of Mucoromycota fungi: Diversity and function of their interactions. In: Timothy James, S.P. (Ed.), The Mycota. Springer Nature.

Vandepol, N., Liber, J., Desirò, A., Na, H., Kennedy, M., Barry, 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 Divers. 104, 267–289.

Waneka, G., Vasquez, Y.M., Bennett, G.M., Sloan, D.B., 2021. Mutational Pressure Drives Differential Genome Conservation in Two Bacterial Endosymbionts of Sap-Feeding Insects. Genome Biol. Evol. https://doi.org/10.1093/gbe/evaa254.

Wang, X., Zhou, H., Chen, H., Jing, X., Zheng, W., Li, R., Sun, T., Liu, J., Fu, J., Huo, L., Li, Y.-Z., Shen, Y., Ding, X., Müller, R., Bian, X., Zhang, Y., 2018. Discovery of recombinases enables genome mining of cryptic biosynthetic gene clusters in Burkholderiales species. PNAS 115, E4255–E24263.

Weete, J., 2012. Fungal Lipid Biochemistry: Distribution and Metabolism. Springer Science & Business Media.

Wernegreen, J.J., 2002. Genome evolution in bacterial endosymbionts of insects. Nat. Rev. Genet. 3, 850–861.

Wernegreen, J.J., 2015. Endosymbiont evolution: predictions from theory and surprises from genomes. Annals of the New York Academy of Sciences 1360, 16–35.

Wernegreen, J.J., Moran, N.A., 1999. Evidence for genetic drift in endosymbionts (Buchnera): analyses of protein-coding genes. Mol. Biol. Evol. 16, 83–97.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the tidyverse. J. Open Source Softw. 4, 1686.

Woolfit, M., Bromham, L., 2003. Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. Mol. Biol. Evol. 20, 1545–1555.

Zhao, H., Nie, Y., Jiang, Y., Wang, S., Zhang, T.-Y., Liu, X.-Y., 2022. Comparative Genomics of Mortierellaceae Provides Insights into Lipid Metabolism: Two Novel Types of Fatty Acid Synthase. J Fungi (Basel) 8. https://doi.org/10.3390/jof8090891.