environmental microbiology

Environmental Microbiology (2017) **00**(00), 00–00 doi:10.1111/1462-2920.13669

Comparative genomics of *Mortierella elongata* and its bacterial endosymbiont *Mycoavidus cysteinexigens*

J. Uehling,[1] A. Gryganskyi,[2] K. Hameed,[1] T. Tschaplinski,[3] P. K. Misztal,[4] S. Wu,[5] A. Desirò,[6] N. Vande Pol,[6] Z. Du,[7] A. Zienkiewicz,[7] K. Zienkiewicz,[7,8] E. Morin,[9] E. Tisserant,[9] R. Splivallo,[10] M. Hainaut,[11] B. Henrissat,[11] R. Ohm,[12] A. Kuo,[13] J. Yan,[13] A. Lipzen,[13] M. Nolan,[13] K. LaButti,[13] K. Barry,[13] A. H. Goldstein,[4] J. Labbé,[3] C. Schadt,[3] G. Tuskan,[3] I. Grigoriev,[13] F. Martin,[9] R. Vilgalys[1] and G. Bonito[6]*

[1]*Department of Biology, Duke University, Durham, NC, 27708, USA.*

[2]*LF Lambert Spawn Company Coatesville, PA, 19320, USA.*

[3]*Biosciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, 37831, USA.*

[4]*University of California Berkeley, Berkeley, CA, 94720, USA.*

[5]*Arizona State University Tempe, AZ, 85281, USA.*

[6]*Plant Soil and Microbial Sciences, Michigan State University, East Lansing, MI, 48824, USA.*

[7]*MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, MI, 48824, USA.*

[8]*Department of Plant Biochemistry, Georg-August University, Göttingen, 37073, Germany.*

[9]*Institut National de la Recherche Agronomique, UMR 1136 INRA-Université de Lorraine 'Interactions Arbres/ Microorganismes', Laboratoire d'excellence ARBRE, INRA-Nancy, Champenoux, 54280, France.*

[10]*Goethe University Frankfurt, Institute for Molecular Biosciences, 60438 Frankfurt, Germany Integrative Fungal Research Cluster (IPF), Frankfurt, 60325, Germany.*

[11]*Architecture et Fonction des Macromolécules Biologiques, CNRS, Aix-Marseille Université, Marseille, 13288, France.*

[12]*Microbiology, Department of Biology, Utrecht University, Utrecht, The Netherlands.*

[13]*Department of Energy, Joint Genome Institute, Oakland, CA, 94598, USA.*

Received 1 November, 2016; revised 5 January, 2017; accepted 7 January, 2017. *For correspondence. E-mail bonito@msu.edu; Tel. (517)884-6958; Fax (517) 353-5174

Summary

Endosymbiosis of bacteria by eukaryotes is a defining feature of cellular evolution. In addition to well-known bacterial origins for mitochondria and chloroplasts, multiple origins of bacterial endosymbiosis are known within the cells of diverse animals, plants and fungi. Early-diverging lineages of terrestrial fungi harbor endosymbiotic bacteria belonging to the Burkholderiaceae. We sequenced the metagenome of the soil-inhabiting fungus *Mortierella elongata* and assembled the complete circular chromosome of its endosymbiont, *Mycoavidus cysteinexigens*, which we place within a lineage of endofungal symbionts that are sister clade to *Burkholderia*. The genome of *M. elongata* strain AG77 features a core set of primary metabolic pathways for degradation of simple carbohydrates and lipid biosynthesis, while the *M. cysteinexigens* (AG77) genome is reduced in size and function. Experiments using antibiotics to cure the endobacterium from the host demonstrate that the fungal host metabolism is highly modulated by presence/absence of *M. cysteinexigens*. Independent comparative phylogenomic analyses of fungal and bacterial genomes are consistent with an ancient origin for *M. elongata* – *M. cysteinexigens* symbiosis, most likely over 350 million years ago and concomitant with the terrestrialization of Earth and diversification of land fungi and plants.

Introduction

Endobacterial symbioses have been reported from several major fungal phyla including Ascomycota and Basidiomycota (Bertaux *et al.*, 2003; Hoffman and Arnold, 2010; Ruiz-Herrera *et al.*, 2015; Arendt et al., 2016) and are especially frequent among early diverging lineages in the Mucoromycota (Bianciotto, 2003; Partida-Martinez and Hertweck, 2005; Sato *et al.*, 2010; Desiro *et al.*, 2013; Naito *et al.*, 2015; Torres-Cortes *et al.*, 2015; Spatafora *et al.*, 2016). Five groups of bacterial endosymbionts of fungi, defined as bacteria living within viable or active fungal cells, have genome sequences available: (1) *Burkholderia rhizoxinica* from *Rhizopus microsporus*

(Lackner *et al*., 2011); (2) *Candidatus* Glomeribacter gigasporarum from *Gigaspora margarita* (Ghignone *et al*., 2012); (3) Mollicutes/*Mycoplasma*-related endosymbionts associated with different genera of Glomeromycotina (Naito *et al*., 2015; Torres-Cortes *et al*., 2015); (4) *Mycoavidus cysteinexigens* from *Mortierella elongata* (Fujimura *et al*., 2014); and (5) *Rhizobium radiobacter* from *Serendipita indica* (Sharma, 2008; Glaeser *et al*., 2015). There has been much study on the evolution and function of bacterial endosymbionts in insects. However, less is known about fungal-endobacteria interactions, the roles of bacterial endosymbionts in fungal evolution and ecology, and the impact of the fungal niche on endosymbiont evolution.

Mortierella is a diverse genus estimated to contain 100–170 species (Nagy *et al*., 2011). These fungi are globally distributed and often dominate environmental fungal communities (Tedersoo *et al*., 2014). Many *Mortierella* species can readily be cultured and have been isolated from plant roots (Bonito *et al*., 2014) and macroalgae (Furbino *et al*., 2014). While *Mortierella* are typically classified ecologically as sugar fungi or soil saprotrophs, their ecology remains poorly understood. Similar to other soil-inhabiting zygomyceteous fungi, *Mortierella* isolates are characterized by rapid growth and multinucleated haploid mycelia with irregular septation, and bidirectional cytoplasmic streaming. *Mortierella* spp. typically exhibit a rosette colony morphology, and a garlic-like odor (Gams, 1977). Given their unique lipid metabolism, some *Mortierella* species are industrially important for dietary supplement production (e.g. omega-3 fatty acids) and biofuel industries (e.g. glycerolipids) (Papanikolaou *et al*., 2007).

Recently, an isolate of *Mortierella elongata* from Japan was found to host betaproteobacteria (Sato *et al*., 2010). Through media supplementation this bacterium was isolated and described as *Mycoavidus cysteinexigens* FMR23-6 I-B1 (Ohshima *et al*., 2016). Together with *Candidatus* Glomeribacter, *Mycoavidus* endosymbionts form a monophyletic lineage that are sister clade to the genus *Burkholderia*. Collectively, we refer to bacteria belonging to the most inclusive monophyletic clade that includes *Candidatus* (*Ca.)* Glomeribacter gigasporarum and *Mycoavidus cysteinexigens* species as the Glomeribacter-Mycoavidus clade. Bacteria belonging to this lineage have previously been visualized within the fungal host cytoplasm through FISH (Desiro *et al*., 2014), Bacteria Counting Kits (Molecular Probes) (Bianciotto *et al*., 2003), and Live-Dead staining kits (Sato *et al*., 2010), but their impact on their host is unknown. Here, we provide the first closed genome for this group of bacteria, the first study on the evolutionary history, and data on the functioning of *M. elongata* and its endosymbiont *M. cysteinexigens*. This was accomplished through metagenome sequencing of *M. elongata* AG77 (host together with *M. cysteinexigens*). Comparative genomic analyses of endobacterial genomes were coupled with experiments comparing metabolic, volatile profiles and growth phenotypes of antibiotic cured and uncured isolates of *Mortierella elongata* AG77.

Results

We screened 30 isolates of *Mortierella* and related fungi isolated from soils and roots of *Populus* (Bonito *et al*., 2016) and detected *Mycoavidus cysteinexigens* in four isolates (~13%) including AG30, AG77, NVP64 and PMI624 (Supporting Information Table S1). We verified the presence of *M. cysteinexigens* in the soil isolate *M. elongata* AG77 and sequenced the genome of this fungus and its bacterium as discussed in detail below (see Supporting Information Fig. S1; Table S1).

The genome of *Mortierella elongata*

The 49 863 165 bp *M. elongata* genome was sequenced to a depth of 112× resulting in 473 contigs and a total of 14 969 predicted gene models with 2467 unique Pfam domains (Supporting Information Table S2). Functional genomic comparisons made through KEGG pathway profile correlations show that *M. elongata* clusters with species in the Mucoromycota (e.g. *Rhizopus*, *Umbelopsis*, *Rhizophagus*) indicative of higher functional genetic similarity between these taxa (Fig. 1, Supporting Information Fig. S1A). Relative to these taxa *Mortierella* is enriched in gene number for most KEGG categories, including fatty acid synthesis and degradation (Supporting Information Table S3A and B). Additionally *M. elongata* is enriched in proteins containing WD 40, FAD, Sel1, protein kinase and Sel1-like domains (Supporting Information Table S3B; Fig. S1B and C).

Fungal carbohydrate and nitrogen metabolism

Gene content and carbon utilization assays indicate *M. elongata* is able to utilize carbon as simple sugars and amino acids. In addition to having a complete set of genes for glycolysis, tricarboxylic acid cycle and protein metabolism, there are several classes of gene families enriched in *M. elongata* compared with other sequenced fungal genomes (Fig. 1). These include genes for glycan biosynthesis and metabolism, 1,4-alpha-glucosidases, as well as amino acid and chitin metabolism, including CAZys belonging to CBM5 and CE4 functional categories (Supporting Information Table S3A). Carbon utilization assays using Biolog plates were consistent with these predicted gene models and demonstrate that *M. elongata* readily utilizes N-acetyl glucosamine (a chitin monomer) as well as the amino acids L-glutamic acid, L, aspartic acid, L-asparagine, L-alanine (Supporting Information Fig. S2). Further, *M. elongata* can utilize simple sugars such as D-glucose, D-trehalose, D-mannose and lipids (tween 20), but not complex organic

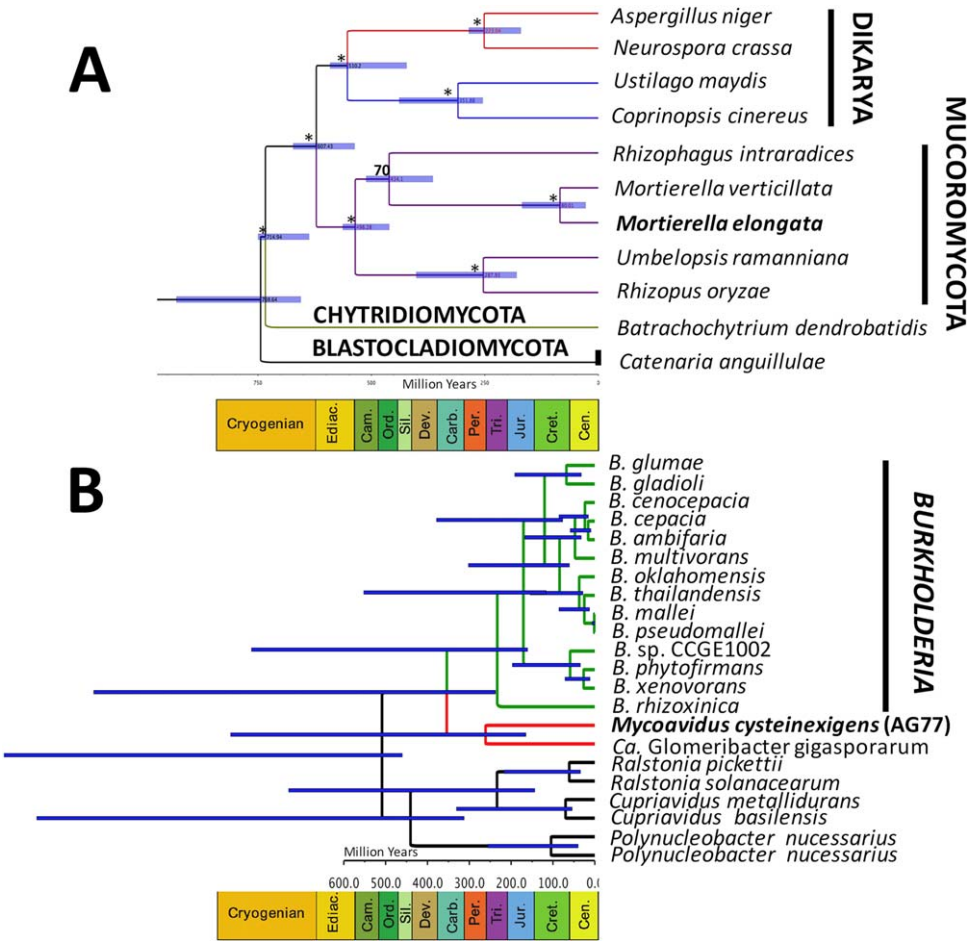

Fig. 1. A. Genome scale divergence time estimation based on BEAST analysis. The maximum clade credibility chronogram estimated in BEAST is shown with nodes placed at the median age (labeled). Node bars (grey) represent the node age 95% highest posterior density (HPD) for nodes receiving at least 50% Bayesian posterior probability. Relationships between *M. elongata* and other fungi with sequenced genomes are shown. Numbers above nodes indicate bootstrap support in RAxML analyses for taxon groupings, with asterisks (*) indicating bootstrap support of 100%. B. Genome scale BEAST analysis illustrating credibility chronogram estimatations and phylogenetic relationships between *Mycoavidus cysteinexigens* (AG77) and other bacteria in the Glomeribacter-Mycoavidus clade having sequenced genomes. Node bars (grey) represent the node age 95% highest posterior density (HPD) for nodes receiving at least 50% Bayesian posterior probability. All indicate bootstrap support of 100%.

polymers such as cellulose or lignin, attributable to a lack of genes encoding carbohydrate-active enzymes (Supporting Information Table S3A).

Fatty acid synthesis in Mortierella elongata

Mortierella species possess the capacity to abundantly produce polyunsaturated fatty acids such as arachidonic acid (Wang *et al.*, 2011). Consistent with these observations, the *M. elongata* AG77 genome contains many copies of genes in fatty acid synthase (FAS) pathways. FAS pathways are divided into types I and II, which are predominant in either animals and fungi or bacteria and plants, respectively (Marrakchi *et al.*, 2002; Schweizer and Hofmann, 2004; Leibundgut *et al.*, 2008; Ploskon *et al.*, 2008). We identified 51 of these genes putatively involved in fatty acid biosynthesis from the *M. elongata* genome including acetyl-CoA carboxylase components, fatty acid synthases, desaturases, elongases, acyl-CoA thioesterase and synthetase (Supporting Information Table S4A). Only one type I FAS gene was identified, which encodes a polypeptide with eight enzymatic domains belonging to a single-chain fungal FAS family, and sharing domain architecture similarity with those in *Coprinopsis cinerea* and *Mortierella alpine* (GenBank Accssion PRJNA211911) (Wang *et al.*, 2011). We also found type I FAS gene orthologues in the *M. verticillata* genome (GenBank accession PRJNA20603) (Supporting Information Table S4A). Intriguingly, we found several type II FASs in the *M. elongata*, *M. alpina*, and *M. verticillata* genomes including genes encoding malonyl-CoA:ACP malonyltransferase, 3-oxoacyl-ACP synthase and 3-Ketoacyl-ACP reductase (Supporting Information Table S4A). While not common, the co-occurrence of type I and type II FASs has been observed in other fungi (Hiltunen *et al.*, 2009; Zhou *et al.*, 2014) (*Saccharomyces cerevisiae*), algae (Hauvermale *et al.*, 2006) (*Nannochloropsis oceanica*) and Apicomplexans (Cai *et al.*, 2005) (*Cryptosporidium parvum*). Taken together, these observations suggest *Mortierella* species, including *M. elongata*, are capable of producing and utilizing diverse fatty acids.

Genome of the endosymbiont Mycoavidus cysteinexigens

The 2 638 116 bp *Mycoavidus cysteinexigens* (AG77) genome was sequenced to a depth of 254× and was

assembled into a single circular chromosome containing 2255 CDS, 6 rRNAs, 41 tRNAs, and a GC content of 49% (Supporting Information Fig. S3; Table S2). Genomes of *Ca*. Glomeribacter gigasporarum (Ghignone *et al*., 2012) and *Mycoavidus cysteinexigens* (FMR23-6 I-B1) (Fujimura *et al*., 2014) have been previously sequenced, however they are fragmented assemblies. We compared the *M. cysteinexigens* (FMR23-6 I-B1) and *M.cysteinexigens* (AG77) genomes and found several rearrangements including inversions and indels (Supporting Information Fig. S4A). When compared with *Ca*. G. gigasporarum, the two *Mycoavidus* genomes have retained or gained unique genes including multiple transcription factors and genes coding for fatty acid metabolism enzymes, while *Ca*. G. gigasporarum is unique in harboring a plasmid (Jargeat *et al*., 2004). Further, both *Mycoavidus* genomes have multiple copies of malate transporters not found in *Ca*. G. gigasporarum. Comparative genomic analyses based upon mapping previously published contigs of *M. cysteinexigens* (FMR23-6 I-B1) and *Ca*. G. gigasporarum (BEG34) genomes to our *M. cysteinexigens* (AG77) unitig are presented in Supporting Information Fig. S4A. Genome comparisons show shared genomic content and synteny between these isolates. Briefly, *M. cysteinexigens* (FMR23-6 I-B1) and *Ca*. G. gigasporarum (BEG34) genomes exhibit between ~80–100% and ~30–85% (Supporting Information Fig. S4B) amino acid similarity, respectively, to the *M. cysteinexigens* (AG77) genome presented here. Compared with free-living *Burkholderia* species, fungal endosymbionts in the Burkholderiaceae appear to have lost many broad functional genes including flagellar biosynthesis, hook-associated, assembly and motor protein genes (see Supporting Information Table S5 for a complete list). Endosymbionts in the Glomeribacter-Mycoavidus clade have lost many other genes. These include gene inactivation and loss of nearly entire alternative glycolytic pathways such as Entner-Douderoff, pyruvate metabolism and the pentose phosphate pathway (Fig. 3 and Supporting Information Fig. S5). Other pathways are characteristically impacted by single gene losses such as in glycolysis, where hexokinase, 6-phosphofructokinase and pyruvate kinase are missing in bacteria within the Glomeribacter-Mycoavidus clade (Supporting Information Fig. S5). In *M. cysteinexigens* (AG77) biosynthetic pathways for the production of numerous amino acids are abridged, including histidine, cysteine, tyrosine, arginine, lysine and asparagine. It is probable that several of these essential amino acids are imported from the host via active transmembrane transporters as discussed below.

Fatty acid metabolism in Mycoavidus cysteinexigens

Although genes are missing for several key enzymes of glycolysis, the *M. cysteinexigens* (AG77) genome has a full suite of genes for the biosynthesis, transport and metabolism of fatty acids, some of which occur in multiple copies. Fatty acid biosynthesis in many bacteria proceeds by the cooperative action of fatty acid synthase complex (type II), a multienzyme protein which catalyzes biosynthesis of fatty acids from acetyl co-A and malonyl coA (Heath and Rock, 1996). The genes involved in fatty acid biosynthesis are in the *Fab* cluster, and we identified all the genes encoding enzymes of bacterial FASII initiation and elongation module in the *M. cysteinexigens* (AG77) genome (Supporting Information Table S4B). Additionally, the *M. cysteinexigens* (AG77) genome possesses two 3-hydroxyacyl-ACP dehydrases encoded by *fabZ* and *fabA* genes, which catalyze the dehydration of various 3-hydroxyacyl-ACPs and the isomerization reaction respectively (Heath and Rock, 1996). The retention of genes contributing to the biosynthesis or manipulation of fatty acids (some in multiple copies) within this highly reduced genome suggests *M. cysteinexigens* (AG77) is able to synthesize and modify fatty acids, potentially for energetic and cellular processes.

Metabolic pathways for fatty acid synthesis and degradation play an important role in bacterial physiology (Cronan, 2003; Yao *et al*., 2012; Yao and Rock, 2013). Transport of extracellular long-chain fatty acids across cell membranes is facilitated by the coordinated action of several genes including acyl-CoA synthetase (*FadD*) and others (Dirusso and Black, 2004). The *M. cysteinexigens* (AG77) genome was found to encode four copies of *FadD*. Despite genome contraction, retention and duplication of this gene set has occurred. Within bacterial cells, fatty acids are degraded via the β-oxidation pathway, a set of enzymes encoded by the *fad* regulon. The complete suite of *Fad* genes (Zhang and Rock, 2016) are present in the *M. cysteinexigens* (AG77) genome (Supporting Information Table S4B) suggesting full capability to degrade host derived saturated and unsaturated fatty acids of various lengths. Given the endocellular nature of this endosymbiont and presence of *FadD* genes, which are absent from related *Burkholderia* spp. such as *B. rhizoxinica*, genome evidence is consistent with the hypothesis that *M. cysteinexigens* (AG77) utilize *M. elongata* derived fatty acids and their breakdown products through β-oxidation.

Secretion systems of Mycoavidus cysteinexigens

The *M. cysteinexigens* (AG77) genome possesses predicted genes for types II, III and IV secretion systems, which may enable translocation of proteins DNA between *M. cysteinexigens* (AG77) and its fungal host. Genomes of close relatives *Burkholderia rhizoxinica* and Ca. *Glomeribacter* also have Type II, III and IV secretion systems (Lackner *et al*., 2011; Ghignone *et al*., 2012). *B. rhizoxinica* and other free-living, pathogenic *Burkholderia* species

contain type III secretion system gene clusters in the *hrp* super family (Lackner *et al*., 2011). In related *Burkholderia* species, these secretion systems are used to translocate effector proteins which manipulate host biology, such as those used by closely related pathogens *B. thailandensis* and *B. pseudomallei* (Stevens *et al*., 2004). In contrast, components of the *M. cysteinexigens* (AG77) type III secretion system (T3SS) share strong homology with gene clusters present in gamma-Proteobacteria such as *Salmonella* and *Yersinia* (Hueck, 1998) rather than the *hrp* systems of closer relatives. *Ca*. G. gigasporarum has a *Salmonella*-like T3SS comparable to that of *M. cysteinexigens* (Ghignone *et al*., 2012). Additionally, in *M. cysteinexigens* there are two predicted Type III Secretion System effector molecules adjacent to the predicted section system components with homology to SseB in *Salmonella* (Nikolaus *et al*., 2001) (Supporting Information Table S6). The alternate T3SS between close relatives indicate a deep divergence between of *Burkholderia* and the Glomeribacter-Mycoavidus clade in concordance with dating analyses presented here, or alternatively, horizontal gene transfer since their divergence.

Predicted secondary metabolite clusters in the Mycoavidus cysteinexigens *(AG77) genome*

The *M. cysteinexigens* (AG77) genome contains six gene clusters putatively involved in secondary metabolite production. One gene cluster has homology with siderophore production, three are predicted as non-ribosomal peptide synthase gene clusters (NRPSs), and two have homology with clusters producing lassopeptides and arylpolyenes respectively (Supporting Information Figs S6–S11; Supporting Information Table S7). The *M. cysteinexigens* (AG77) genome contains several transporters for NRPS and secondary metabolites, presumably enabling transport of these products from endosymbiont to fungus. In Ascomycota fungi, secondary metabolite gene clusters are used for the production of stress-induced compounds including antibiotics (Spatafora and Bushley, 2015), such as genes which can be often involved in plant-fungal and fungal-bacterial interactions (Partida-Martinez and Hertweck, 2007; Soanes and Richards, 2014). It is notable that *M. elongata* has fully functional primary metabolism but lacks secondary metabolite gene clusters (including those for antibiotic production), a genome trait that is shared with other early diverging terrestrial fungi (Tisserant *et al*., 2013). In contrast, while the *M. cysteinexigens* (AG77) genome lacks many genes necessary for primary metabolism it retains several secondary metabolite and NRPS gene clusters. Both fungal host and *M. cysteinexigens* (AG77) genomes possess predicted transmembrane transporters capable of importing/exporting secondary metabolites. It is plausible that the export of antimicrobial compounds is important to the chemical ecology of *Mortierella* and its *M. cysteinexigens* endosymbiont.

Predicted transporters in the Mycoavidus cysteinexigens *(AG77) genome*

Over fifty predicted genes involved in trans-membrane substrate transport occur in the *M. cysteinexigens* (AG77) genome. These include ABC transporters for methionine, histidine, proline and general amino acids. Other predicted transporters include those for nitrate, nitrite, nucleosides, dipeptides, phosphate and phosphonate (Fig. 3). Based on evidence for specific transporters, several primary metabolite products may be imported from the host including glycerol-3-phosphate, malate and other C4-dicarboxylates (Fig. 3). The *M. cysteinexigens* (AG77) genome also includes transporters for vitamin B7 (biotin) and vitamin B12 (cobalamin), which catalyze fatty acid and amino acid metabolism (Fig. 3). Transporters were also detected for uptake of Zn^{2+}, Mg^{2+}, $Fe^{2+,3+}$, Na^{+} and K^{+} ions (Fig. 3). Based on its genome annotation, *M. cysteinexigens* (AG77) may possess the capacity to make and export siderophores, lipopolysaccharides, drugs and metabolites via resistance-nodulation-cell division superfamily (RND) efflux pumps and drug metabolite transporters (DMTs) (Fig. 3).

Phylogenomic analysis of Mortierella elongata

Phylogenetic relationships of the earliest diverging fungal lineages have been challenging to determine, especially for fungi previously classified as Zygomycota (James *et al*., 2006; Stajich *et al*., 2009). To better understand the evolutionary history of *M. elongata* we used a genome dataset that included 494 single copy orthologous genes from 11 taxa. Phylogenomic analyses support placement of *M. elongata* in a monophyletic clade with other early-diverging terrestrial fungi belonging to the Mortierellomycotina, Mucoromycotina and Glomeromycotina (Fig. 4A; Supporting Information Fig. S12A) – following the classification of Spatafora *et al*. (2016). Cluster analysis of shared KEGG functional proteins also suggests strong functional similarity among these fungi (Fig. 1). Fungal fossils attributed to arbuscular mycorrhizal spores have been dated to be over 460 million years old (Redecker and Graham, 2000; Taylor *et al*., 2015), and the divergence of related *Glomeromycotina–Ca*. Glomeribacter symbiosis was previously established to be at least 400 million years ago (Mondo *et al*., 2012). We used genome data to estimate the divergence between Glomeromycotina, Mortierellomycotina and other fungi. Our divergence time estimates indicate these fungi diverged from other early diverging lineages during the Cambrian and Ordivician (460-558 MYA) and

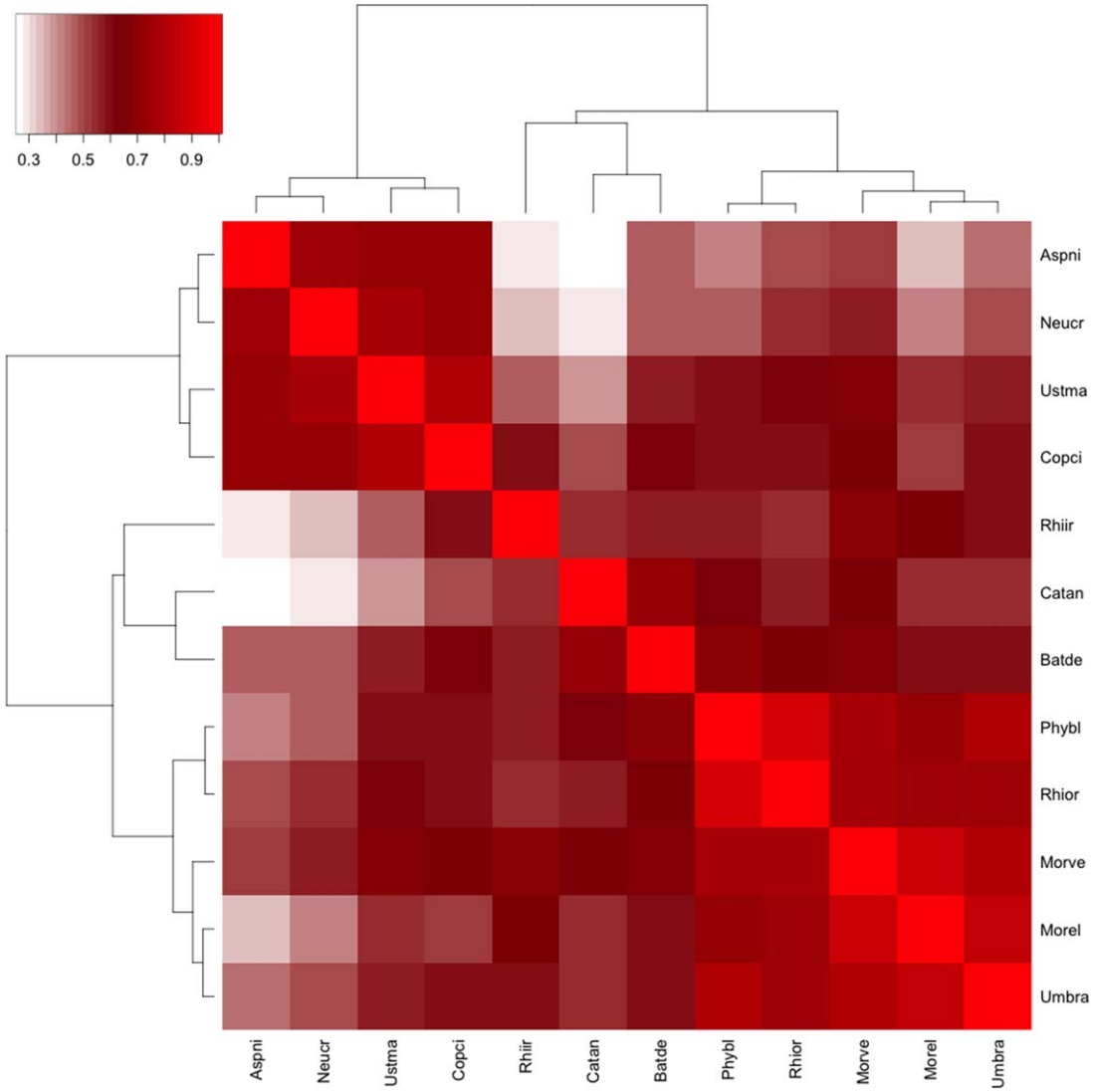

Fig. 2. Correlation of KEGG pathway profiles between 12 fungal species from the early diverging Fungi (including M. elongata, black arrow) and Dikarya. Pearson correlation distance matrix was calculated based on presence/absence profile of protein-coding genes assign to each KEGG orthology for each species. Colors are coded from dark red representing high correlation to light red representing low correlation. (Aspni: Aspergillus niger; Neucr: Neurospora crassa; Ustma: Ustilago maydis; Copci: Coprinopsis cinerea; Rhiir: Rhizophagus irregularis; Catan: Catenaria anguillulae; Batde: Batradochytridium dendrobaditis; Phybl: Rhior: Rhizopus oryzae; Rhizopus oryzae; Morve: Mortierella verticillata; Mortierella elongata; Umbra: Umbelopsis ramanniana.

from each other between the Devonian-Cambrian periods (358-508 MYA) (Fig. 4A).

Phylogenomic analysis of the Mycoavidus cysteinexigens

Phylogenomic analyses using 20 taxa and a core set of 10 single copy orthologous genes are consistent with 16S rDNA data with high bootstrap (100%) support for the evolutionary placement of the Glomeribacter-Mycoavidus clade as sister to *Burkholderia* (β-proteobacteria; Burkholderiaceae) (Fig. 4B; Supporting Information Fig. S12B). Analyses also strongly support the independent origin for the endosymbiotic bacterium *B. rhizoxinica*, which is placed within the genus *Burkholderia*. Gene content analyses revealed similar patterns of gene loss in each lineage of endobacteria within the Glomeribacter-Mycoavidus clade (Fig. 2) relative to *Burkholderia* representatives that include obligate symbionts and free-living bacteria, indicative of convergent evolution. BEAST analyses based on a core set of 10 genes across the Burkholderiaceae estimate median divergence times between *Burkholderia* and Glomeribacter-Mycoavidus clade at 350 million years ago (MYA), and between *Ca*. Glomeribacter and *Mycoavidus*

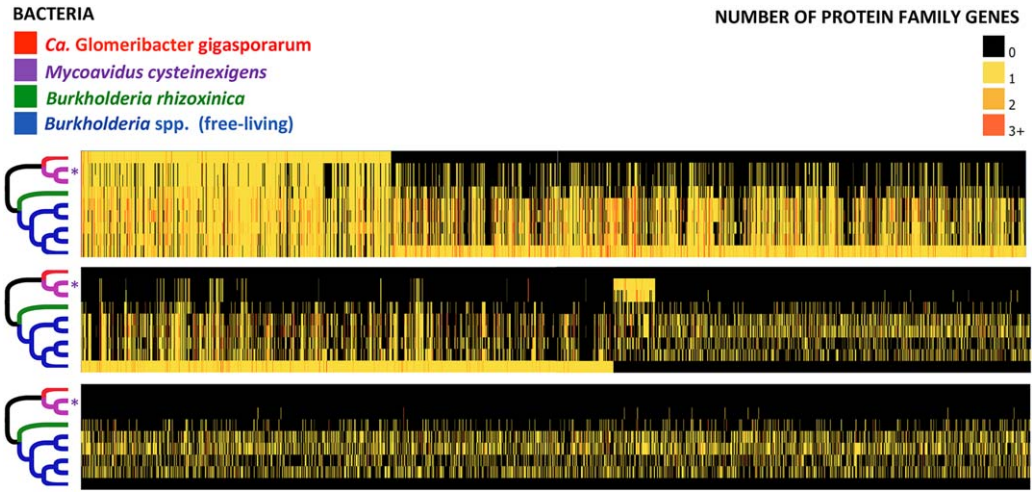

Fig. 3. Heat map showing numbers of 5634 protein family genes (FigFAMS) observed across eight different Burkholdariaceae genomes. Bacterial genomes are indicated by the phylogenetic tree on the left, from top: *Mycoavidus cysteinexigens* (AG77) of *Mortierella elongata* (AG77), indicated by asterisk*, *Candidatus Glomeribacter* gigasporarum BEG34, *Burkholderia rhizoxinica* HKI454, *B. thailandensis* E264, *B. mallei* SAVP1, *B. mallei* ATCC 23344, *B. pseudomallei* 1026b, *B. pseudomallei* 1710b. Colors on phylogenetic tree indicated obligate (red) or facultative (pink) endosymbiotic bacteria, and free-living *Burkholderia* spp. (blue), (including the animal pathogen *B. mallei*). Gene sequences belonging to a single FIGfam are isofunctional homologues inferred to have same functional role, with similarity over at least 70% of the length of the protein coding sequences. FigFAM heat map was generated using PATRIC (Wattam et al. 2014).

at least 260 MYA, but likely earlier (Fig. 4B). Although the phylogenomic divergence dating analysis has large confidence intervals, owing to poor calibration points for bacteria, these results are consistent with the hypothesis that both endosymbiont and fungus were diverging from close relatives during the same geologic periods.

Functional analyses of Mycoavidus cysteinexigens *on* Mortierella elongata *fungal host colony*

Transmission electron microscopy. We have obtained another isolate of *M. elongata* (NVP64) that contains the same endobacterium as AG77 (*M. cysteinexigens* – Supporting Information Table S1). This isolate of *M. elongata* was more amenable to sporulation and imaging. We used transmission electron microscopy in order to confirm the location of *M. cysteinexigens* within the hyphae of *M. elongata* and to describe its morphology (Fig. 5). Numerous *M. cysteinexigens* bacterial cells were observed within the *Mortierella* mycelium: they were rod-shaped, 300-450 $\times$ 600-900 nm in size, with a jagged and layered, Gram-negative type-like cell wall, and a cytoplasm rich in ribosomes consistent with previous reports (Sato *et al*., 2010) (Fig. 5A and B). In contrast to the phylogenetically related *Ca*. G. gigasporarum no clear evidence of a membrane of fungal origin surrounding the endobacteria was observed. However, a tangled complex of membranes was visible around and between the bacterial cells (Fig. 5B). Endobacteria occurred in groups constituted by many individuals often in proximity of large lipid bodies that filled the hyphal space throughout its width (Fig. 5A). A difference in lipid body quantity and dimension was visually evident between uncured and cured strains; in particular, a lower number of large lipid bodies were observed in the uncured (Fig. 5A) and numerous small lipid droplets were observed with a scattered distribution in the cured strain cytoplasm (Fig. 5C).

• *Effect of* Mycoavidus *removal on the growth of* Mortierella elongata. *M. elongata* AG77 was cured of its *M. cysteinexigens* endosymbiont using a panel of antibiotics used previously to cure the fungus *R. microsporus* of *B. rhizoxinica* endobacteria (Partida-Martinez and Hertweck, 2005). Clearing of *M. elongata* was confirmed through 16S rDNA PCR assays using general and specific primers (Supporting Information Table S8). Higher growth rates were observed in the cleared isolate ($p < 0.05$, Students *t*-test) (Fig. 6A). In comparison with the cured *M. elongata* isolate, which had more symmetric and rosette-like colonies with well-developed aerial hyphae typical of *Mortierella*, uncured isolate AG77 grew more slowly and produced less aerial hyphae (Fig. 6B).

Vertical transmission of *M. cysteinexigens* (AG77) has occurred in isolates grown in pure culture (removed from soils and healthy plant roots) and maintained over a five-year interval of growth and agar plate transfer in the lab. Interestingly, zygospores and sporangiospores are uncommon in both uncured and cured isolates of *Mortierella elongata*, particularly *M. elongata* AG77.

• *Effects of* Mycoavidus cysteinexigens *on fungal metabolism.* To study the effect of *Mycoavidus* on its host metabolism we grew *M. elongata* in potato dextrose broth (PDB) and PDB supplemented with peptone, and

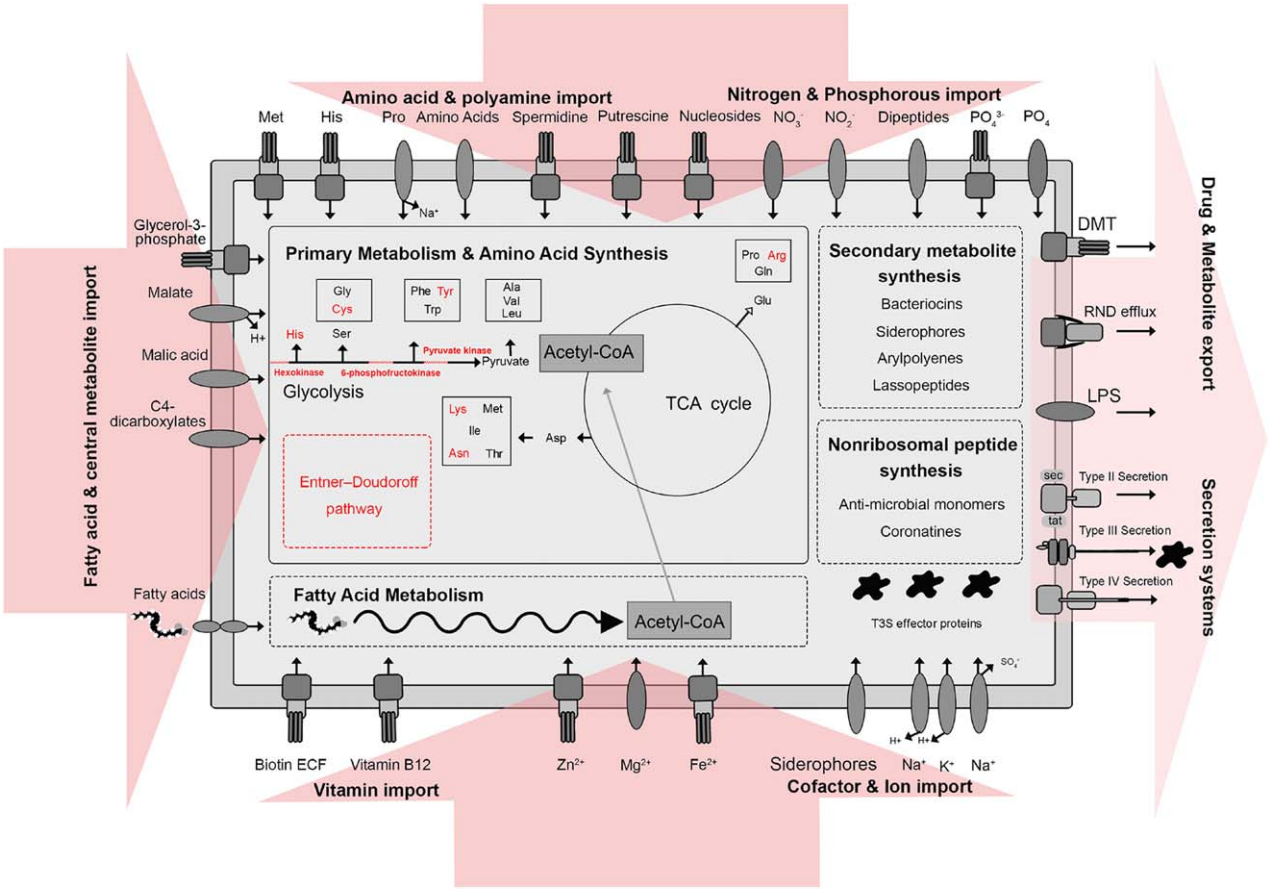

Fig. 4. Schematic illustrating *Mycoavidus cysteinexigens* (AG77) genome content. Gene loss for amino acids and enzymes involved in primary metabolism indicate abridged pathways and limited metabolism in M. cysteinexigens. Larger red arrows indicate directionality of products into and out of the *M. cysteinexigens* endosymbiont. Rounded transporters represent transmembrane transporters annotated as having specific cargo in the genome. Squared transporters represent ABC transporters. Symporter and antiporters are annotated with their cargo and ions. Secondary metabolite and nonribosomal peptide synthetase gene products reflect predictions of AntiSMASH and homology with the Norine database.

contrasted the differential metabolite profiles of cured and uncured isolates with gas chromatography-mass spectrometry (GCMS). All observations were based on three replicates, and fold change significance was evaluated using Student *t*-test, $p < 0.05$ (Fig. 6C, red bars indicate significant observations). Fold changes are reported as a ratio of a given metabolite in the cleared strain, over the uncleared strain such that 1× fold indicates no chance, >1× fold indicated an increase in the cleared strain and <1× fold change indicates a decrease in the cleared strain. Overall, we observed declines in storage carbohydrates, organic acids and amino acids (0–1× fold), and accumulation of fatty acids (2–6× fold) in the cured isolate (Fig. 6C). We inferred from analyses of *Mycoavidus* gene retention and gene class enrichment that fungal derived fatty acids fuel *M. cysteinexigens* metabolism. Consistent with this hypothesis we observed an accumulation of the long chain saturated fatty acids and branched fatty acids (2–6× fold, Fig. 6C) in the cured isolate. These products included iso-myristic acid (2.43× fold), palmitic acid (2.04× fold), and tetracosanoic acid (3.82× fold), 11-eicosenoic acid (3.35× fold), stearic acid (3.59× fold) and arachidic acid (6.74× fold) (Fig. 6C). In addition to a broad-scale increase in many fatty acids in the cured isolate, several related fatty acids and metabolites decreased in the *Mycoavidus* uncured isolate, including dodecanoic acid (0.47× fold), isopentadecanoic acid (0.53× fold), gamma-linolenic acid (0.59× fold) and monoolein (0.53× fold) (Supporting Information Table S9), perhaps reflective of shifting fatty acid utilization. In contrast, we observed mostly decreases in storage carbohydrates and amino acids in the cured isolate. For example, primary metabolites and amino acids that significantly decreased in the cured isolate included maltose (0.09× fold), fructose (0.65× fold), lactic acid (0.51× fold) and citric acid (0.55× fold), isoleucine (0.25× fold), valine (0.27× fold) and leucine (0.27× fold). In summary, curing of *Mycoavidus* from *M. elongata* resulted in a significantly altered fungal colony

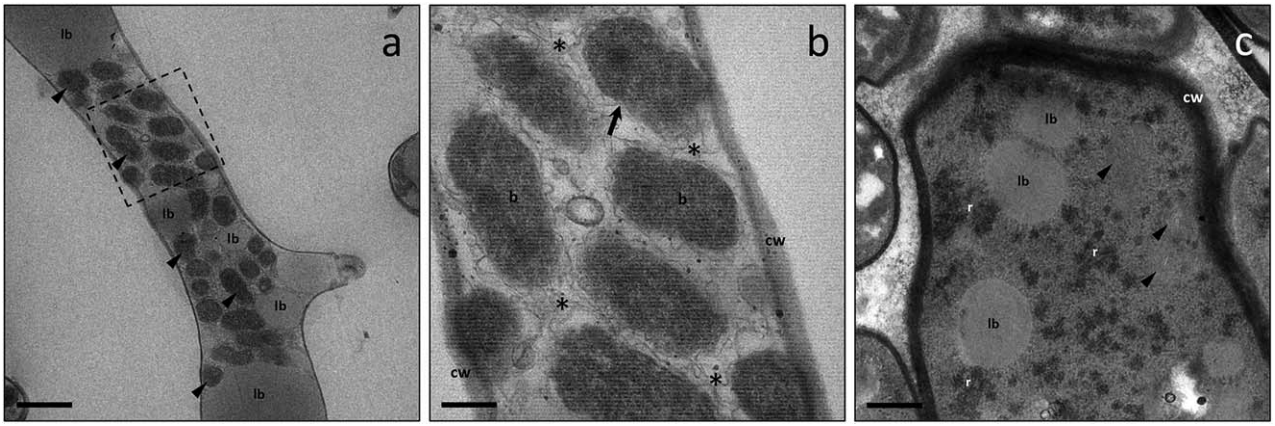

Fig. 5. Transmission electron microscopy of wild type (a, b) and cured (c) strains of *Mortierella elongata* NVP64.
A. A group of clustered rod-shaped *Mycoavidus* cells (arrowhead) within the fungal mycelium in proximity of several, usually large, lipid bodies (lb).
B. Magnification of square from panel (a): *Mycoavidus* cells (NVP64) (b) have a jagged cell wall and are surrounded by a tangled complex of membranes (asterisk). Endobacteria are engaged in cell division, as suggested by the central constriction (arrow).
C. The cured strain showed a dense cytoplasm, rich in mitochondria (arrowhead), ribosomal aggregates (r) and small lipid bodies (lb). Fungal cell wall (cw). Bars: (a) 1 μm; (b) 0.26 μm; (c) 0.37 μm.

metabolism that involved an accumulation of fatty acid products in concert with declines in storage carbohydrates, organic acids and nitrogenous metabolites. Whether these shifts are due directly to endosymbiont metabolism dynamics, or are part of a fungal response to endosymbiont presence remains to be tested. Taken together with growth rate data, declines in carbohydrates and amino acids in the cured isolate relative to uncured are likely attributed to the greater energy demand by *M. cysteinexigens* (AG77).

• *Effects of* Mycoavidus cysteinexigens *(AG77) on the volatile profile of* Mortierella elongata*.* Cured isolates of *M. elongata* also exhibit change in colony odor, which has been variously described as similar to garlic or 'wet dog' (Gams, 1977). To quantify these volatile emissions profiles of cured and uncured isolates were compared using proton-transfer mass spectrometry (PTR-MS). Cured *M. elongata* emitted fewer VOCs (Fig. 6D) consistent with fatty acid breakdown products such as butyric and crotonic acid and their esters, which (normalized to CO_2) increased 2- to 20-fold (Fig. 6D; Supporting Information Fig. S13). Curing also resulted in ~30% lower respiration, suggesting either *M. elongata*, *Mycoavidus* or both display altered physiological functioning and activity when in symbiosis compared with *M. elongata* alone (Supporting Information Fig. S14).

Volatile profiles were also compared with solid phase microextraction-GCMS, providing an alternative snapshot of emitted products. *Mortierella elongata* hosting *M. cysteinexigens* produced higher concentrations of alcohols, aldehydes, ketones, furans and one unidentified volatile (Fig. 6E; Supporting Information Table S10). These induced volatiles contained metabolites with 8 carbon atoms (octen-3-ol, 3-octanone, 1-octen-3-one), which are generally regarded as a class of fungal hormones (Chitarra *et al*., 2005). In contrast, one sesquiterpenoid (tentatively identified as such based on the known characteristic masses m/z 164, 149, 109) is repressed by the presence of *Mycoavidus* within the fungal mycelium (Supporting Information Table S10). We hypothesize these volatiles influence microbial interactions and chemical ecology in the soil and plant rhizosphere.

Discussion

Plant-associated fungi belonging to the phylum Mucoromycota (Spatafora *et al*., 2016) are known to harbor endosymbiotic bacteria belonging to the Glomeribacter-Mycoavidus clade (Bianciotto *et al*., 1996; Sato *et al*., 2010; Mondo *et al*., 2012). Our phylogenomic and divergence-time estimates indicate that *Mortierellomycotina* and *Glomeromycotina* diverged from a common ancestor between 358 and 508 MYA. This is consistent with other published estimates for this lineage of fungi (Malloch *et al*., 1980; Chang *et al*., 2015). It is notable that the endobacterium *M. cysteinexigens* (AG77) (within *Mortierella*) and *Ca*. Glomeribacter (within Glomeromycotina*)* are vertically transmitted and share a common ancestor (this study and Lumini *et al*., 2007). Divergence dating of bacteria is still quite challenging given the lack of bacteria in fossil records and high levels of horizontal gene transfer. Consequently, our efforts to estimate median divergence dates for the Burkholderaceae resulted in wide confidence intervals and distribution skewed to the left indicating all lineages may be older than MYA dates indicate. Even so, our estimates of 350 MYA for the Mycoavidus-Glomeribacter

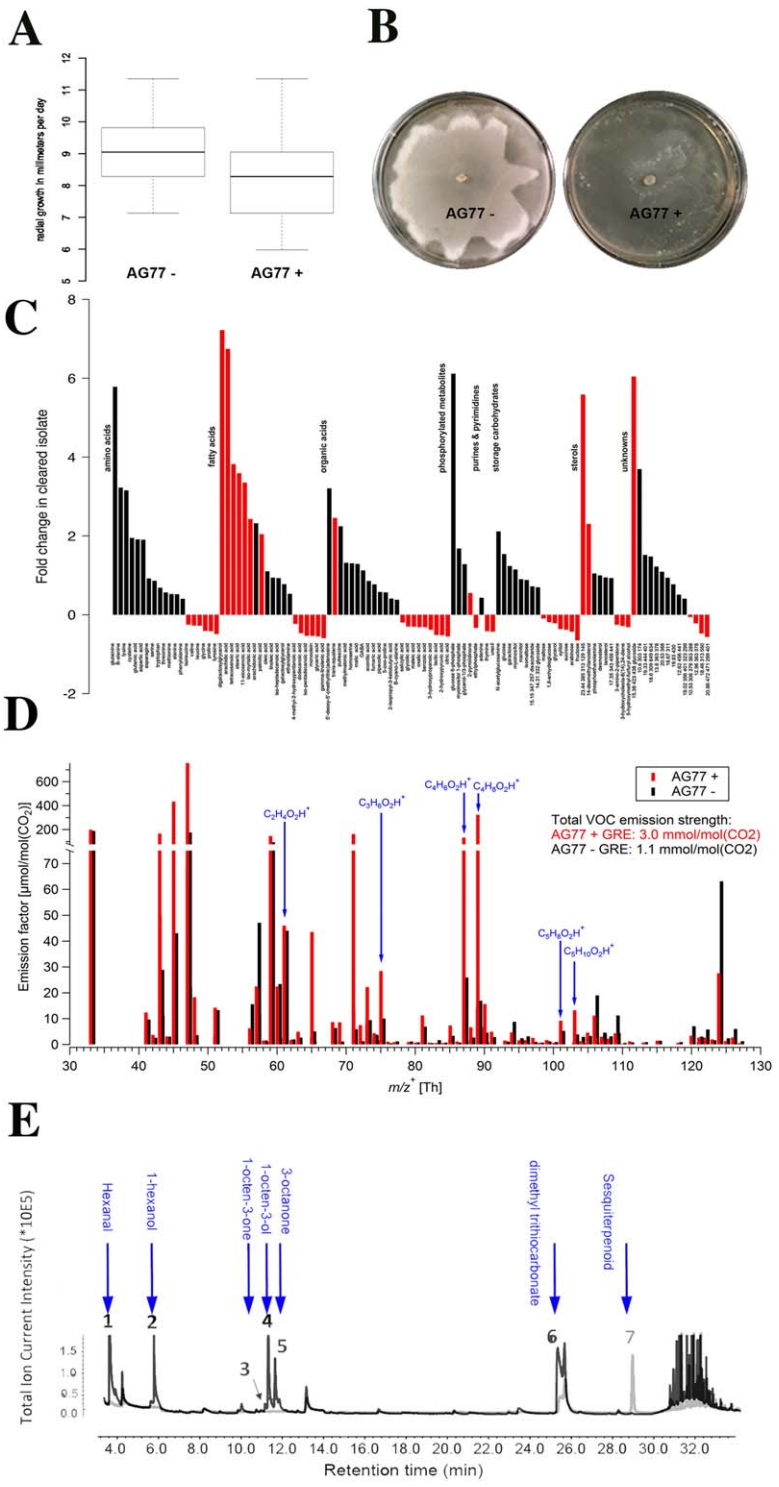

Fig. 6. Morphological and metabolomics shifts in *M. elongata* with the presence and absence of *M. cysteinexigens* (AG77).
A. Radial fungal growth per day in millimeters on malt extract agar. For each fungal strain the cleared strains are those subjected to antibiotic passaging. Stars indicate a significant shift in growth rate between wild type and cleared strains.
B. M. elongata on Malt Extract liquid media with (left) and without (right) *M. cysteinexigens* (AG77).
C. Fold increase of metabolites in cleared strains, by category. Red bars indicate significance across replicates based on $p < 0.05$, Students *t* test.
D. Fingerprints of mVOC emission factors of cleared and uncleared strain AG77. Chemical formulas were derived from measured exact mass and are shown here for the peaks corresponding to short-chain fatty acids. Gray trace has been offset by 0.3 Th for clarity. Note: the chemical formulas may represent more than one structure. For example, *m/z* 61.028 ($C_2H_4O_2H_+$) is the sum of acetic acid and acetate, *m/z* 75.043 ($C_3H_6O_2H_+$) is the sum of propionic acid and propionate, *m/z* 87.043 ($C_4H_6O_2H_+$) is the sum of crotonic acid, isocrotonic acid, crotonate and isocrotonate, *m/z* 89.059 ($C_4H_8O_2H_+$) is the sum of butyric acid, isobutyric acid, butyrate and isobutyrate, *m/z* 101.059 ($C_5H_8O_2H_+$) is the sum of methylcrotonate and methylisocrotonate, and *m/z* 103.074 ($C_5H_{10}O_2H_+$) is the sum of methylbutyrate and methylisobutyrate. E. Fingerprints of volatiles detected by GC-MS. Gray trace has been offset for clarity. Blue arrows denote products that match known database standards.

lineage indicate an ancient origin of bacterial endosymbiosis in the Mucoromycota. Previous researchers report 400 million years as the age of the Glomeromycota-*Ca*. Gomeribacter symbiosis based on parametric co-phylogeny models and the date of the oldest arbuscular mycorrhizal fossil (Mondo *et al.*, 2012). Our results based on genomic data, additional taxon sampling and divergence estimates provide a more realistic range and confidence in estimating the origin of fungal-bacterial endosymbioses between these biological lineages.

It is evident that endobacteria belonging to the Glomeribacter-Mycoavidus clade have evolved into unique

niches based on genome content, gene loss and cultivability compared with each other and related free-living bacteria. We suggest that this divergence may be also due to selective forces imposed upon them by their host physiology. For instance, *Mortierella* (saprotrophic, easily cultured) and fungi in the Glomeromycotina (obligate biotrophic, fastidious to culture) have distinct ecological roles and diverged long ago in evolutionary time. Fungal hosts in these lineages likely impose different selection pressures upon their endosymbionts, potentially explaining some of the intrageneric genomic and functional differences presented here and presented by others (Ghignone *et al*., 2012; Fujimura *et al*., 2014).

The higher growth rate and vigor of *M. elongata* cured of *M. cysteinexigens* indicates there is a fitness cost to on the host to harbor endosymbionts under the conditions tested, including increased respiration and fungal fatty acid catabolism of the mycelial colony. In contrast, previous studies of *Ca*. G. gigasporarum in *Gigaspora margarita* showed that cleared fungal strains exhibited decreased growth and plant host fitness associated with oxidative stress amelioration (Lumini *et al*., 2007; Salvioli *et al*., 2015; Vannini *et al*., 2016). It is less clear why a fungus would maintain an energetically costly endosymbiont for hundreds of millions of years. In many symbioses, novel functionality is gained by utilizing gene repertoires of multiple organisms, leading to subtle or context dependent symbiotic benefits. For instance, the closely related endosymbiont *B. rhizoxinica* provides its host *R. microsporus* bacterially derived toxins that facilitate plant tissue invasion by the fungal host offsetting the physiological cost of hosting endobacteria (Lackner *et al*., 2009). Proteomics analysis of the same cured and uncured isolates of *Mortierella* (AG77) presented here show that the metabolism of the fungal host and bacterium are independent but closely intertwined (Li *et al*., 2016). Although the types of trade-offs involved in the *Mortierella-Mycoavidus* symbiosis are not fully clear yet, the presence of secondary metabolite gene clusters within the *M. cysteinexigens* (AG77) genome offers intriguing potential tradeoff. Does *Mycoavidus* improve the competitive interaction of its fungal host with other microbes in the rhizosphere and soil environment? Co-culturing cleared and wild type fungal strains in a community context through soil and plant bioassays may be useful for testing this hypothesis.

Our culture studies demonstrate that the metabolism of *M. elongata* is significantly altered in the absence of *Mycoavidus*, which likely underpins their interactions. Microbial genome sequencing, comparative metabolomics, growth assays and volatile quantification assays of cleared fungal strains support this hypothesis. Specifically, both host-endosymbiont genomes and functional data we present highlight the importance of fungal fatty acid metabolism in this long co-evolution, which appears to have been coopted by the endosymbiont *M. cysteinexigens*. Altered fatty acid metabolism has been previously implicated in fungal-bacterial interaction dynamics (Deveau *et al*., 2007) and fungal endosymbiosis (Salvioli *et al*., 2010). It is possible that the application of an antibiotic cocktail could adversely alter fungal phenotypes. However, it is the strain that endured high antibiotic treatments that are consistently more fit and productive, counter to what would be expected if the antibiotics were deleterious to the fungal host. To unambiguously test these hypotheses future studies involving endosymbiont re-infection will be necessary.

It is a well-documented phenomenon in human and plant microbiome research that host-bacterial physiology and metabolism are intertwined at multiple levels (Turnbaugh *et al*., 2009, Koenig *et al*., 2011, Muegge *et al*., 2011, Berendsen *et al*., 2012). Here, we provide evidence for an ancient fungal-bacterial endosymbiosis in the Mucoromycota based upon host-endosymbiont genome sequencing and analysis. We show that the bacterial endosymbiont *M. cysteinexigens* appears to catabolize fungal fatty acids, and that host metabolism and volatile profiles of the fungus *M. elongata are* significantly altered by the absence of *M. cysteinexigens*. The implications of this interaction on the ecology of *Mortierella* remain unknown, but given the widespread distribution of *Mortierella* in soils they may be ecological relevant. As more bacterial endosymbionts of fungi are discovered their study will provide a unique perspective for understanding eukaryotic host-symbiont interactions and fungal evolution.

Methods

Isolation/nucleic acid extractions

Mortierella elongata AG77 was isolated on 1% potato dextrose agar (PDA) from fresh soils collected in Duke Forest NC, USA following a dilution plating technique. Additional strains of *Mortierella* and zygomycetous fungi were isolated from Duke Forest NC (AG77), *Populus* roots from along the Yadkin River in NC, and in agricultural soils of Michigan, USA (Bonito *et al*., 2016). We amplified ITS and LSU rDNA and 16S rDNA (for bacterial screening) from isolates and Sanger sequenced them. Sequences were aligned and compared through phylogenetic analyses, as described below (Supporting Information Fig. S15). To prepare fungal mycelium for genome sequencing the isolate AG77 was grown in liquid potato dextrose broth for five days. Fungal mycelium was collected and washed twice in sterile water before nucleic acid extraction. DNA was extracted with CTAB 2x following the DNA chloroform extraction technique. RNA was extracted from *Mortierella elongata* mycelium with the Qiagen RNeasy plant mini kit (Qiagen, Valencia, CA, USA) following manufacturer's instructions to aid in fungal genome annotation. Total RNA quality was quantified using a 2100 bioanalyser (Agilent, Santa Clara, CA, USA) and Qubit 2.0 fluorometer (Invitrogen, Grand Island, NY, USA) as specified in the supplier's protocols. RNA showed high integrity

(RIN > 8.0) and well-defined peaks for 18S and 28S rRNAs on an electropherogram (RNA 6000 Nano LabChip, Agilent), and was used for cDNA library construction. Genome isolate AG77 is deposited at the CBS-KNAW Fungal Biodiversity Center under accession number 137287.

Fungal and bacterial genome sequencing

For the *Mortierella elongata* genome, Illumina fragment and long mate pair libraries were sequenced, assembled with AllPathsLG, and then further improved with Pacific Biosciences data. For the fragment library, 100 ng of genomic DNA was sheared using the Covaris E220 (Covaris) and sized selected using SPRI beads (Beckman Coulter). The DNA fragments were treated with end repair, A-tailing, and ligation of Illumina compatible adapters (IDT, Inc.) using the KAPA-Illumina library creation kit (KAPA biosystems). LFPE (ligation-free paired end) mate pair fragments were generated using the 5500 SOLiD Mate-Paired Library Construction Kit (SOLiD®). 15 μg of genomic DNA was sheared using the Covaris g-TUBETM (Covaris) and gel size selected for 4 kb. The sheared DNA was end repaired, and ligated with biotinylated internal linkers. The DNA was circularized using intra-molecular hybridization of the internal linkers. The circularized DNA was treated with plasmid safe to remove non-circularized products. The circularized DNA was nick translated and treated with T7 exonuclease and S1 nuclease to generate fragments containing internal linkers with genomic tags on each end. The mate pair fragments were A-tailed and purified using Strepavidin bead selection (Invitrogen). The purified fragments were ligated with Illumina adaptors and amplified using 8 cycles of PCR with Illumina primers (Illumina) to generate the final library.

Both libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v3, and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flowcell was performed on the Illumina HiSeq2000 sequencer using a TruSeq SBS sequencing kit, v3, following a 2x100 and 2x150 indexed run recipe for LFPE and fragments respectively. Genomic reads from two libraries were filtered and assembled with AllPathsLG (Gnerre *et al.*, 2011). PacBio data were used to fill gaps in Illumina genome assembly. For this, 5 μg of DNA was sheared to 10 kb using the g-TUBE^{TM} (Covaris). The sheared DNA was treated with DNA damage repair mix followed by end repair and ligation of SMRT adapters using the PacBio SMRTbell Template Prep Kit (PacBio). PacBio sequencing primer was annealed to the SMRTbell template library and Version XL sequencing polymerase was bound to them. The prepared SMRTbell template libraries were then sequenced on a Pacific Biosciences RSII sequencer using Version C2 chemistry and running 1 × 120 min movies per SMRT Cell. The data was QC filtered for artifact/process contamination and subsequently assembled together with Illumina assembly using PBJelly. The closed circular bacterial chromosome was completed using a mixture of Illumina and PacBio reads. For PacBio, raw data was aligned to a draft assembly of the symbiont to create a list of reads to keep and run with RS_HGAP_Assembly_3.3 with smrtpipe version 2.3.0. Overlapping contig ends were trimmed. Since the main chromosome was in a close circular contig other small contigs were excluded. Assembly statistics are summarized in Supporting Information Table S2.

Annotation

The fungal genome was annotated using the JGI Annotation pipeline and made available via JGI fungal genome portal MycoCosm (Grigoriev *et al.*, 2013) (jgi.doe.gov/fungi). The bacterial genome was annotated using the JGI Microbial Genome Annotation Papeline (Huntemann *et al.*, 2015). Annotation statistics is summarized in Supporting Information Table S2. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession LYLZ00000000. Raw sequence data and analysis files are available from the NCBI-BioProject (PRJNA196039) and JGI Mycocosm website (http://genome.jgi.doe.gov/Morel2/Morel2.home.html) respectively. The assembled *M. cysteinexigens* (AG77) genome is publically available through the PATRIC website (www.patricbrc.org/) under the genome ID 224135.3.

Transcriptome sequencing for scaffolding genome assembly, gene prediction and validation

Transcriptomes were sequenced using two different protocols: with rRNA depletion and polyA selection, each using 1 μg of total RNA. For the former, rRNA was was removed using Ribo-Zero^{TM} rRNA Removal Kit (Bacteria) (Epicentre). For the latter, mRNA was purified using magnetic beads containing poly-T oligos. Stranded cDNA libraries were generated using the Illumina Truseq Stranded RNA LT kit. RNA was fragmented and reverse transcribed using random hexamers and SSII (Invitrogen) followed by second strand synthesis. The fragmented cDNA was treated with end-pair, A-tailing, adapter ligation, and 10 cycles of PCR. All libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified library was then multiplexed into a pool of 4 libraries, and the pool was then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v4, and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flowcell was performed on the Illumina HiSeq2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a 2x150 indexed run recipe. RNA-seq data for each genome were *de novo* assembled into consensus sequences using Rnnotator (v. 2.5.6 or later) (Martin *et al.*, 2010).

Comparative bacterial genomics analyses

KEGG pathways were manually examined using the PATRIC web genome browser, based on the RASTtk annotation (Brettin *et al.*, 2015). Shared genome content with relatives was analysed using the heat map function of the comparative pathway map available through the PATRIC workbench (Wattam *et al.*, 2014). Antimicrobials and secondary metabolites were predicted based on comparisons of the JGI annotation files, RAST, AntiSMASH pipeline (Medema *et al.*, 2011), and

Norine (Caboche *et al.*, 2008). Gene order and homology with relatives was determined using generous and BLAST. Figures were made using PhotoShop (Adobe, San Jose, CA, USA).

Phylogenetic analyses

To identify orthologous clusters of genes having exactly one member in each of the organisms genome sequences from selected fungal taxa were subject to an OrthoMCL (Li *et al.*, 2003) analysis using protein sequences. Sequences for each cluster were aligned independently using mafft. Poorly aligned regions were removed through with the software program Gblocks (Talavera and Castresana, 2007). Resulting alignments were concatenated and PartitionFinderProtein (Lanfear *et al.*, 2012) was used to determine best substitution models and partitioning schemes for the concatenated alignment. Phylogenetic analyses based on maximum likelihood as calculated through RAxML (Stamatakis, 2014) through the CIPRES (Miller *et al.*, 2010) portal to calculate most likely phylogenetic trees. Thousand bootstrap replicates were run to determine supported nodes. Phylogenomic alignments were deposited in (TreeBase #18877).

For fatty acid and lipid genomic comparisons within the *Mortierella* protein sequences of *M. elongata* genes involved in fatty acid and glycerolipid synthesis were downloaded from MycoCosm (http://genome.jgi.doe.gov/pages/search-for-genes.jsf?organism = Morel2). Protein sequences of *M. alpina* B6842 (PRJNA211911) and *M. verticillata* NRRL 6337 (PRJNA13353) were deduced from genome annotations downloaded from Genbank. To identify potential homologues, we performed a protein-protein BLAST analysis (BLAST+ 2.2.31) using default parameters.

Divergence dating of *Mortierella elongata*

Time to the most common ancestor (TMRCA) were estimated using BEAST v1.8.1(Drummond *et al.*, 2012). The analyses were run for 100 million iterations and sampled every 1000 iterations. Four independent chains were run and the posterior probability was inspected manually using Tracer v1.6(Rambaut *et al.*, 2014) to ensure convergence. The first 20% of the samples were discarded as burn-in and the 95% Highest Posterior Density region (HPD) were calculated using Tracer.

The TMRCA were estimated using the following parameter setting. Each dataset were fitted with LG amino acid substitution model (Le *et al.*, 2008). The Yule speciation model and uncorrelated lognormal-distributed relaxed clock model were employed. A user-defined starting tree was used given the size and complexity of the dataset. This starting tree was generated using RAxML with 1000 bootstraps and with *Catenaria anguillae* as outgroup. This agrees with our knowledge of evolutionary trends in basal fungal lineages (Chang *et al.*, 2015), therefore the topology was fixed for this analysis. Two informative priors were used, the prior distribution for the root is a truncated uniform distribution between 760 MYA and 1060 MYA (Chang *et al.*, 2015). The prior distribution for *Rhizophagus irregularis* was a truncated uniform distribution between 460 MYA and 600 MYA (Simon *et al.*, 1993). All other priors are BEAST default priors. We attempted to estimate the TMRCA between *Rhizophagus irregularis* and nearest living relative sampled using a 95% highest posterior density region between 460 MYA and 495 MYA.

Divergence dating of the glomeribacter-mycoavidus clade

Time to the most common ancestor (TMRCA) were estimated using BEAST v1.8.1 (Drummond *et al.*, 2012). There were 22 taxa and there were no missing taxa in any partitions. Ten partitions of single copy core genes were combined together to estimate the TMRCA on the same phylogeny. The LG amino acid substitution model (Le *et al.*, 2008) coupled with Yule speciation model and uncorrelated lognormal distributed relaxed clock were used in the analyses. The prior for the root was an uniform distribution between 460 MYA (estimated split between *Mortierella* and *Rhizophagus*) and 2500 MYA (the great oxidation event), with BEAST default settings for all other priors. The analyses started with a randomly generated phylogeny and were run for 100 million iterations and sampled every 10000 iterations. Four independent chains were run and the posterior probability was inspected manually using Tracer v1.6 (Rambaut *et al.*, 2014) to ensure convergence. The first 10% of the samples were discarded as burn-in and the 95% Highest Posterior Density region (HPD) were calculated using Tracer.

Computational analysis of gene family evolution (CAFE)

Multigene families were predicted from 164 143 predicted proteins found in the 12 genomes using the MCL algorithm with an inflation parameter set to 3.0. As a result, 4921 protein families were identified. Multigene families were analysed for evolutionary changes in protein family size using the CAFE program. The program uses a random birth and death process to model gene gain and loss across a user specified tree structure. The distribution of family sizes generated under the random model provides a basis for assessing the significance of the observed family size differences among taxa (p-value 0.001). CAFE estimates for each branch in the tree whether a protein family has not changed, has expanded or contracted. The phylogenetic tree used is the one constructed as described above.

MCL and pFAM Computational analysis of gene family evolution (CAFE) analyses

We performed phylogenomic analysis using 12 genomes (taxa listed in Fig. 4), and identified 387 gene families with only one gene per species by clustering protein sequences using FastOrtho (http://enews.patricbrc.org/fastortho/). Each family was then aligned with MAFFT 7.221, and ambiguous regions (containing gaps and poorly aligned) were eliminated and single-gene alignments were concatenated with Gblocks (Talavera and Castresana, 2007). We achieved a Maximum Likelihood inference for our phylogenomic dataset with RAxML 7.7.2 using the standard algorithm, the PROTGAMMAWAG model of sequence evolution and 1000 bootstrap replicates.

Multigene families were predicted from 164 143 predicted proteins found in the 12 genomes using the MCL algorithm with an inflation parameter set to 3.0. As a result, 4921 protein families were identified. Multigene families were analysed for

evolutionary changes in protein family size using the CAFE program (De Bie *et al.*, 2006). The program uses a random birth and death process to model gene gain and loss across a user specified phylogenetic tree depicted in (Supporting Information Fig. S1B), constructed according method in paragraph Phylogeny. The distribution of family sizes generated under the random model provides a basis for assessing the significance of the observed family size differences among taxa (*p*-value 0.001) (Supporting Information Fig. S1B). CAFE estimates for each branch in the tree whether a protein family has not changed, has expanded or contracted.

Comparative bacterial genome analyses

Genomes sequences *Mycoavidus cysteinexigens*, (WGS RefSeq NZ_BBOF00000000.1) and *Candidatus* Glomeribacter gigasporarum Beg34 (WGS RefSeq accession NZ_CAFB00000000.1) were downloaded from Joint Genome Institute's GenBank. The dotplot matrix (Supporting Information Fig. S4A) was generated using the Blast DotPlot viewing option. The *M. cysteinexigens* and *Ca*. Glomeribacter gigasporarum were aligned to the *Mycoavidus cysteinexigens* (AG77) genome and compared using the RAST based annotations and SEED Viewer version 2.0 (Overbeek *et al.*, 2005). Amino acid similarity was compared with blastP, using the complete *Mycoavidus cysteinexigens* (AG77) genome as a reference (Supporting Information Fig. S4B).

Antibiotic clearing of endobacteria from *Mortierella*

Endobacteria were cleared from AG77 using antibiotics as previously described (Partida-Martinez and Hertweck, 2007). Cultures were cycled fifteen times between solid and liquid media (one week growth for each transfer) using 60 μg/ml Kanamycin, Streptomycin and Chloramphenicol, and 100 μg/ml Ciprofloxacin. Passaging and comparative growth analyses were performed in duplicate.

Bacterial detection and primer design

Fungal endosymbiotic bacteria were monitored in cured and non-cured isolates and verified using primer pair 8F/1492R (Reysenbach *et al.*, 1992; Baker *et al.*, 2003) and the newly designed *Glomeribacter* specific 16S primer at 60° C annealing temperature (Supporting Information Fig. S15). Primers new to this study (Supporting Information Table S5) were designed in Primer Blast (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and tested for specificity against close relatives of *Ca*. Glomeribacter within the Burkholderiaceae (data not shown).

Growth assay and analyses

Cleared and uncleared fungal isolates were plated on 1.5% MEA and incubated at 25°C. Agar plugs of uniform size inoculated with fungal cultures were placed at the center of single plates and radial growth measured in centimeters in each of four quadrats every 24 h. For each cleared and uncleared isolate pair, three technical replicates were measured. Students *T*-test was performed to detect statistical significance ($P < 0.05$) growth between cleared and uncleared isolates (Fig. 6).

Metabolomics

Fungal mycelium of cleared and uncleared representatives of each isolate were grown for 7 days at 28°C with constant agitation in 1.5% potato dextrose broth and 1.5% potato dextrose broth with 2g of peptone added (Difco Laboratories, Detroit, MI, USA). Fresh mycelia were filtered and washed with sterile water, frozen in liquid nitrogen, and freeze-dried. For metabolomic profiling, 50 mg of freeze dried mycelia were ground with a micro-Wiley mill and twice extracted with 2.5 ml 80% ethanol overnight, and 0.5-ml aliquots were dried in a nitrogen stream. Sorbitol was added (to achieve 15 ng/μL injected) before extraction as an internal standard to correct for differences in extraction efficiency due to subsequent differences in changes in sample volume during heating. Dried extracts were silylated as described previously (Li *et al.*, 2012; Tschaplinski *et al.*, 2012). After 2 days, 1-μL aliquots were injected into an Agilent Technologies Inc. (Santa Clara, CA) 5975C inert XL gas chromatograph-mass spectrometer, configured and operated as described earlier(Li *et al.*, 2012; Tschaplinski *et al.*, 2012). Metabolite peaks were extracted using a key selected ion, characteristic m/z fragment, rather than the total ion chromatogram, to minimize integrating co-eluting metabolites. The extracted peaks of known metabolites were scaled back up to the total ion current using predetermined scaling factors. Peaks were quantified by area integration and the concentrations were normalized to the quantity of the internal standard (sorbitol) recovered, amount of sample extracted, derivatized, and injected. Metabolites of interest were quantified using a large user-created database (>2300 spectra) of mass spectral electron impact ionization (EI) fragmentation patterns of trimethylsilyl-derivatized compounds and the Wiley Registry 8th Edition combined with NIST 05 mass spectral database.

Volatile profiling with proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF-MS)

A Proton-Transfer-Reaction Time-of-Flight Mass Spectrometer (PTR-TOF-MS - model 8000, Ionicon, Austria) (Jordan *et al.*, 2009) was used to measure volatile organic compound (VOC) emissions. The instrument consists of three parts: (1) the ion source region where primary ions (hydronium ions) are generated from water vapor via ion-molecule reactions in the plasma discharge; (2) the reaction chamber region (also called drift tube) where the hydronium ions softly collide with the VOC in the air introduced to form protonated ions; and (3) the detection region consisting of a TOF-MS (Supporting Information Fig. S16). The technique allows for estimating the concentrations of organic ions over a broad m/z range (1.000-500.000 Th) from proton reaction theory and measured detector transmission (referred to as the 'transmission approach') with reaction rate constants (k) that are known or approximated (e.g. for unidentified compounds). A combination of calibration techniques using authentic standards and the transmission approach results in better than 8% accuracy. For estimating concentration of unidentified compounds it is necessary to use the default k (2×10^{-9} cm^3 s^{-1}), which leads to ~30% uncertainty if the molecule is stable and not fragmenting significantly. Further information on quantification of concentrations and formula identification are described in (Graus *et al.*, 2010; Holzinger *et al.*, 2010). Analyses using

PTR-ToF data for mVOCs are different from conventional techniques such as gas chromatography in that the counting statistics are extremely high due to the real time character of data acquisition. At 1 s time resolution, a half hour sampling period provides thousands of samples and the precision of the signal variability due to instrumental noise is thus dependent on signal averaging time. This leads to high confidence in the results as long as the differences exceed natural variations of the signal by three times standard deviation. The fold enhancements in short chain fatty acids exceeded by orders of magnitude the standard deviation of the variability within the samples so are assumed significant. Due to evident differences, additional statistical analysis such as ANOVA was not found to be necessary.

Volatile profiling with GCMS

Mortierella elongata cured and uncured isolates were grown as four independent cultures per isolate at room temperature in 100 ml Erlenmeyer flask containing 30 ml of malt extract broth (1% weight/vol., Difco Laboratories GmbH, Heidelberg, Germany) in distilled water, pH adjusted to 7.0. After 12 days, the mycelia were transferred to new 100 ml Erlenmeyer flask containing 30 ml of potato dextrose broth (26.5 g l^{-1} from Carl Roth GmbH, Karlsruhe, Germany), yeast extract (4.0 g l^{-1} from AppliChem GmbH, Darmstadt, Germany), pH adjusted to 7.0. The mycelia were harvested after five days and a subsample of 1.20 ± 0.05 g was transferred to a 20 ml airtight vial equipped with PTFE a septum for volatile profiling.

Volatiles from $n = 4$ samples per isolate were profiled using a GCMS equipped with the Solid Phase Micro Extraction (SPME) autosampler as described previously (Molinier *et al.*, 2015) with slight modifications. Essentially, volatiles were extracted from the head-space of the vials for 15 min at 50°C with a DBV/CARD/PDMS 1.0 cm SPME fibre from Supelco (Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany). The GC method consisted in an initial oven temperature: 40°C (hold for 3 min), followed by a ramp at 1.5°C min^{-1} to 80°C; then a ramp at 80°C min^{-1} to 250°C (hold 7.21 min) - total run time: 39.00 min. Chromatograms were processed for peak realignment using TagFinder (Luedemann *et al.*, 2008) and the identification of volatiles was confirmed with authentic standards for 3-methyl-1-butanol, 2-methyl-1-butanol, 1-pentanol, 1-hexanol, 1-octen-3-ol, 3-hydroxy-2-butanone, 1-octen-3-one, and 3-octanone (purchased from Sigma-Aldrich Chemie GmbH, Taufkirchen, Germany). Other volatiles were tentatively identified based on mass spectral data and Kovats retention indices (NIST 2011 Mass Spectral Library, National Institute of Standards and Technology, Gaithersburg, MD, USA). Statistical differences in the concentration of volatiles between isolates were assessed using the non-parametric Kruskal-Wallis tests performed in R (https://www.r-project.org/).

Transmission electron microscopy

Wild type and cured strains of *Mortierella elongata* NVP64 were maintained in malt extract agar (MEA) plates at room temperature. In order to avoid fungal hyphae to grow inside the agar and make the subsequent hyphal collection easier, the mycelium was grown on an autoclaved cellophane sheet laid on the surface of the medium. After 30 days, small mycelium fragments (2×2 mm) were excised from the older part of the cultures in the central portion of the plate, and fixed in 50 mM cacodylate buffer (pH 7.2) containing 2.5% glutaraldehyde for 1 h at room temperature and afterwards overnight at 4°C. The fragments were then rinsed three times with cacodylate buffer and post-fixed in OsO_4 for 1 h. After rinsing twice with the cacodylate buffer, the fragments were progressively dehydrated in an ethanol series and then incubate twice in absolute acetone (Hoch, 1986). The fungal samples were then embedded in fresh Epon-Araldite resin and polymerase for 36 h at 60°C. Semi-thin (1 μm) and ultra-thin sections (70 nm) were cut and processed as described in (Desirò *et al.*, 2016). Ultrastructural analyses were performed by using a JEOL100 CXII transmission electron microscope.

Acknowledgements

This research was sponsored by the Genomic Science Program, U.S. Department of Energy, Office of Science - Biological and Environmental Research as part of the Plant Microbe Interfaces Scientific Focus Area (http://pmi.ornl.gov) at Oak Ridge National Laboratory. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05- 00OR22725. Genome sequencing and annotation was supported by the DOE Joint Genome Institute by the JGI Community sequencing program project 570 'Metatranscriptomics of Soil Forest Ecosystems'. The work conducted by the US Department of Energy's Joint Genome Institute is supported by the Office of Science of the US Department of Energy under Contract DE-AC02-05CH11231. GMB, NVP and AD are grateful to MSU's AgBioResearch for helping to support this research. KZ was supported by a Marie Curie International Outgoing Fellowship within the EU 7th Framework Program. Measurements of VOC by AHG and PKM were supported by the Sloan Foundation Microbiology of the Built Environment program. Research in the laboratory of FM is funded by the Laboratory of Excellence Advanced Research on the Biology of Tree and Forest Ecosystems (ARBRE; grant ANR-11-LABX-0002-01). The authors thank Joseph Spatafora and members of the the ZyGoLife consortium (zygolife.org) for discussions and access to genomic data of zygomycetes for comparative analyses.

Conflict of interest statement

The authors declare there are no conflicts of interest.

References

Arendt, K.R., Hockett, K.L., Araldi-Brondolo, S.J., Baltrus, D.A., and Elizabeth Arnold, A. (2016) Isolation of endohyphal bacteria from foliar ascomycota and in vitro establishment of their symbiotic associations. *Appl Environ Microbiol* **82:** 2943–2949.

Baker, G., Smith, J.J., and Cowan, D.A. (2003) Review and re-analysis of domain-specific 16S primers. *J Microbiol Methods* **55:** 541–555.

Berendsen, R.L., Pieterse, C.M., and Bakker, P.A. (2012) The rhizosphere microbiome and plant health. *Trends Plant Sci* **17:** 478–486.

Bertaux, J., Schmid, M., Prevost-Boure, N.C., Churin, J.L., Hartmann, A., Garbaye, J., and Frey-Klett, P. (2003) In situ identification of intracellular bacteria related to Paenibacillus spp. in the mycelium of the ectomycorrhizal fungus laccaria bicolor S238N. *Appl Environ Microbiol* **69:** 4243–4248.

Bianciotto, V., Bandi, C., Minerdi, D., Sironi, M., Titchy, H.V., and BonFante, P. (1996) An obligately endosymbiotic mycorrhizal fungus itself harbors obligately intracellular bacteria. *Appl Environ Microbiol* **62:** 3005–3010.

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

Bonito, G., Reynolds, H., Robeson, M.S., Nelson, J., Hodkinson, B.P., Tuskan, G., *et al.* (2014) Plant host and soil origin influence fungal and bacterial assemblages in the roots of woody plants. *Mol Ecol* **23:** 3356–3370.

Bonito, G., Hameed, K., Ventura, K., Krishnan, R., Schadt, J.C., and Vilgalys, R. (2016) Isolating a functionally relevant guild of fungi from the root microbiome of Populus. *Fungal Ecol.* **22:** 35–42.

Brettin, T., Davis, J.J., Disz, T., Edwards, R.A., Gerdes, S., Olsen, G.J., *et al.* (2015) RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* **5:**8365.

Caboche, S., Pupin, M., Leclère, V., Fontaine, A., Jacques, P., and Kucherov, G. (2008) NORINE: a database of nonribosomal peptides. *Nucleic Acids Res* **36:** D326–D331.

Cai, X., Herschap, D., and Zhu, G. (2005) Functional characterization of an evolutionarily distinct phosphopantetheinyl transferase in the apicomplexan Cryptosporidium parvum. *Eukaryot Cell* **4:** 1211–1220.

Chang, Y., Wang, S., Sekimoto, S., Aerts, A.L., Choi, C., Clum, A., *et al.* (2015) Phylogenomic analyses indicate that early fungi evolved digesting cell walls of algal ancestors of land plants. *Genome Biol Evol* **7:** 1590–1601.

Chitarra, G.S., Abee, T., Rombouts, F.M., and Dijksterhuis, J. (2005) 1-Octen-3-ol inhibits conidia germination of Penicillium paneum despite of mild effects on membrane permeability, respiration, intracellular pH, and changes the protein composition. *FEMS Microbiol Ecol* **54:** 67–75.

Cronan, J.E. (2003) Bacterial membrane lipids: where do we stand? *Annu Rev Microbiol* **57:** 203–224.

De Bie, T., Cristianini, N., Demuth, J.P., and Hahn, M.W. (2006) CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* **22:** 1269–1271.

Desiro, A., Naumann, M., Epis, S., Novero, M., Bandi, C., Genre, A., and Bonfante, P. (2013) Mollicutes-related endobacteria thrive inside liverwort-associated arbuscular mycorrhizal fungi. *Environ Microbiol* **15:** 822–836.

Desiro, A., Salvioli, A., Ngonkeu, E.L., Mondo, S.J., Epis, S., Faccio, A., *et al.* (2014) Detection of a novel intracellular microbiome hosted in arbuscular mycorrhizal fungi. *ISME J* **8:** 257–270.

Desirò, A., Salvioli, A., and Bonfante, P. (2016) Investigating the endobacteria which thrive in arbuscular mycorrhizal fungi. *Methods in Molecular Biology* (Vol. **1399).** Springer. Clifton, NJ, pp. 29–53.

Deveau, A., Palin, B., Delaruelle, C., Peter, M., Kohler, A., Pierrat, J.C., *et al.* (2007) The mycorrhiza helper Pseudomonas fluorescens BBc6R8 has a specific priming effect on the growth, morphology and gene expression of the ectomycorrhizal fungus Laccaria bicolor S238N. *New Phytol* **175:** 743–755.

Redecker, D., R.K., and Graham, L.E. (2000) Glomalean fungi from the Ordovician. *Science* **289:** 1884–1885.

Dirusso, C.C., and Black, P.N. (2004) Bacterial long chain fatty acid transport: gateway to a fatty acid-responsive signaling system. *J Biol Chem* **279:** 49563–49566.

Drummond, A.J., Suchard, M.A., Xie, D., and Rambaut, A. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29:** 1969–1973.

Fujimura, R., Nishimura, A., Ohshima, S., Sato, Y., Nishizawa, T., Oshima, K., *et al.* (2014) Draft genome sequence of the betaproteobacterial endosymbiont associated with the fungus Mortierella elongata FMR23-6. *Genome Announc* **2.**

Furbino, L.E., Godinho, V.M., Santiago, I.F., Pellizari, F.M., Alves, T.M., Zani, C.L., *et al.* (2014) Diversity patterns, ecology and biological activities of fungal communities associated with the endemic macroalgae across the Antarctic Peninsula. *Microb Ecol* **67:** 775–787.

Gams, W. (1977) A key to the species of Mortierella. *Persoonia* **9:** 381–391.

Ghignone, S., Salvioli, A., Anca, I., Lumini, E., Ortu, G., Petiti, L., *et al.* (2012) The genome of the obligate endobacterium of an AM fungus reveals an interphylum network of nutritional interactions. *ISME J* **6:** 136–145.

Glaeser, S.P., Imani Alabid, J., Guo, I., Kumar, H., Kämpfer, N., Hardt, P., *et al.* (2015) Non-pathogenic Rhizobium radiobacter F4 deploys plant beneficial activity independent of its host Piriformospora indica. *ISME J.* **10:** 871–884.

Gnerre, S., MacCallum, I., Przybylski, D., Ribeiro, F.J., Burton, J.N., Walker, B.J., *et al.* (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc Natl Acad Sci* **108:** 1513–1518.

Graus, M., Muller, M., and Hansel, A. (2010) High resolution PTR-TOF: quantification and formula confirmation of VOC in real time. *J Am Soc Mass Spectrom* **21:** 1037–1044.

Grigoriev, I.V., R., Nikitin, S., Haridas, A., Kuo, R., Ohm, R., Otillar, R., *et al.* (2013) MycoCosm portal: gearing up for 1000 fungal genomes. *Nucleic Acids Res* **42:** D699–704.

Hauvermale, A., K.J., Rosenzweig, B., Guerra, J., Diltz, S., and Metz, J. (2006) Fatty acid production in Schizochytrium sp.- Involvement of a polyunsaturated fatty acid synthase and a type I fatty acid synthase. *Lipids* **41:** 739–747.

Heath, R.J., and Rock, C.O. (1996) Roles of the FabA and FabZ β-hydroxyacyl-acyl carrier protein dehydratases in *Escherichia coli* fatty acid biosynthesis. *J Biol Chem* **271:** 27795–27801.

Hiltunen, J.K., Schonauer, M.S., Autio, K.J., Mittelmeier, T.M., Kastaniotis, A.J., and Dieckmann, C.L. (2009) Mitochondrial fatty acid synthesis type II: more than just fatty acids. *J Biol Chem* **284:** 9011–9015.

Hoch, H.C. (1986) Freeze-substitution of fungi. *Ultrastructure Techniques for Microorganisms*. Springer, US. Plenum Press, New York. pp. 183–212.

Hoffman, M.T., and Arnold, A.E. (2010) Diverse bacteria inhabit living hyphae of phylogenetically diverse fungal endophytes. *Appl Environ Microbiol* **76:** 4063–4075.

Holzinger, R., Kasper-Giebl, A., Staudinger, M., Schauer, G., and Röckmann, T. (2010) Analysis of the chemical composition of organic aerosol at the Mt. Sonnblick observatory using a novel high mass resolution thermal-desorption proton-transfer-reaction mass-spectrometer (hr-TD-PTR-MS). *Atmos Chem Phys* **10:** 10111–10128.

Hueck, C.J. (1998) Type III protein secretion systems in bacterial pathogens of animals and plants. *Microbiol Mol Biol Rev* **62:** 379–433.

Huntemann, M., Ivanova, N.N., Mavromatis, K., Tripp, H.J., Paez-Espino, D., Palaniappan, K.*, et al.* (2015) The standard operating procedure of the DOE-JGI Microbial Genome Annotation Pipeline (MGAP v. 4). *Stand Genomic Sci* **10:** 1.

James, T.Y., Kauff, F., Schoch, C.L., Matheny, P.B., Hofstetter, V., Cox, C.J.*, et al.* (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* **443:** 818–822.

Jargeat, P., Cosseau, C., Ola'h, B., Jauneau, A., Bonfante, P., Batut, J., and Becard, G. (2004) Isolation, free-living capacities, and genome structure of "Candidatus Glomeribacter gigasporarum," the endocellular bacterium of the mycorrhizal fungus Gigaspora margarita. *J Bacteriol* **186:** 6876–6884.

Jordan, A., Haidacher, S., Hanel, G., Hartungen, E., Märk, L., Seehauser, H.*, et al.* (2009) A high resolution and high sensitivity proton-transfer-reaction time-of-flight mass spectrometer (PTR-TOF-MS). *Int J Mass Spectrom* **286:** 122–128.

Koenig, J.E., Spor, A., Scalfone, N., Fricker, A.D., Stombaugh, J., Knight, R., Angenent, L.T., and Ley, R.E. (2011) Succession of microbial consortia in the developing infant gut microbiome. *Proc Natl Acad Sci* **108:** 4578–4585.

Lackner, G., Moebius, N., and Hertweck, C. (2011) Endofungal bacterium controls its host by an hrp type III secretion system. *ISME J* **5:** 252–261.

Lackner, G., Partida-Martinez, L.P., and Hertweck, C. (2009) Endofungal bacteria as producers of mycotoxins. *Trends Microbiol* **17:** 570–576.

Lackner, G., Nadine, M., Partida-Martinez, L.P., Boland, S., and Hertweck, C. (2011) Evolution of an endofungal lifestyle: Deductions from the *Burkholderia rhizoxinica* genome. *BMC Genomics* **12:** 210.

Lanfear, R., Calcott, B., Ho, S.Y., and Guindon, S. (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Mol Biol Evol* **29:** 1695–1701.

Le, S.Q., Lartillot, N., and Gascuel, O. (2008) Phylogenetic mixture models for proteins. *Philos Trans R Soc Lond B: Biol Sci* **363:** 3965–3976.

Leibundgut, M., Maier, T., Jenni, S., and Ban, N. (2008) The multienzyme architecture of eukaryotic fatty acid synthases. *Curr Opin Struct Biol* **18:** 714–725.

Li, L., Stoeckert, C.J., and Roos, D.S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res* **13:** 2178–2189.

Li, Y., Tschaplinski, T.J., Engle, N.L., Hamilton, C.Y., Rodriguez, M., Liao, J.C., *et al.* (2012) Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. *Biotechnol Biofuels* **155C:** 50–56.

Li, Z., Yao, Q., Dearth, S.P., Entler, M.R., Castro Gonzalez, H.F., Uehling, J.K.*, et al.* (2016) Integrated proteomics and metabolomics suggests symbiotic metabolism and multimodal regulation in a fungal-endobacterial system. *Environ Microbiol*. doi: 10.1111/1462-2920.13605

Luedemann, A., Strassburg, K., Erban, A., and Kopka, J. (2008) TagFinder for the quantitative analysis of gas chromatography—mass spectrometry (GC-MS)-based metabolite profiling experiments. *Bioinformatics* **24:** 732–737.

Lumini, E., Bianciotto, V., Jargeat, P., Novero, M., Salvioli, A., Faccio, A.*, et al.* (2007) Presymbiotic growth and sporal morphology are affected in the arbuscular mycorrhizal fungus *Gigaspora margarita* cured of its endobacteria. *Cell Microbiol* **9:** 1716–1729.

Malloch, D., Pirozynski, K., and Raven, P. (1980) Ecological and evolutionary significance of mycorrhizal symbioses in vascular plants (a review). *Proc Natl Acad Sci* **77:** 2113–2118.

Marrakchi, H., Zhang, Y.M., and Rock, C.O. (2002) Mechanistic diversity and regulation of Type II fatty acid synthesis. *Biochem Soc Trans* **30:** 1050–1055.

Martin, J., Bruno, V.M., Fang, Z., Meng, X., Blow, M., Zhang, T.*, et al.* (2010) Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. *BMC Genomics* **11:** 663.

Medema, M.H., Blin, K., Cimermancic, P., de Jager, V., Zakrzewski, P., Fischbach, M.A.*, et al.* (2011) AntisMash: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res* **39:** W339–W346.

Miller, M.A., W., Pfeiffer, and T., Schwartz. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Pages 1-8 *in* Gateway Computing Environments Workshop (GCE), 2010. New Orleans, LA. IEEE.

Molinier, V., Murat, C., Frochot, H., Wipf, D., and Splivallo, R. (2015) Fine-scale spatial genetic structure analysis of the black truffle Tuber aestivum and its link to aroma variability. *Environ Microbiol* **17:** 3039–3050.

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

Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., González, A., Fontana, L.*, et al.* (2011) Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332:** 970–974.

Nagy, L.G., Petkovits, T., Kovács, G.M., Voigt, K., Vágvölgyi, C., and Papp, T. (2011) Where is the unseen fungal diversity hidden? A study of *Mortierella* reveals a large contribution of reference collections to the identification of fungal environmental sequences. *New Phytologist* **191:** 789–794.

Naito, M., Morton, J.B., and Pawlowska, T.E. (2015) Minimal genomes of mycoplasma-related endobacteria are plastic and contain host-derived genes for sustained life within Glomeromycota. *Proc Natl Acad Sci U S A* **112:** 7791–7796.

Nikolaus, T., Deiwick, J., Rappl, C., Freeman, J.A., Schröder, W., Miller, S.I., and Hensel, M. (2001) SseBCD proteins are secreted by the type III secretion system of Salmonella pathogenicity island 2 and function as a translocon. *J Bacteriol* **183:** 6036–6045.

Ohshima, S., Sato, Y., Fujimura, R., Takashima, Y., Hamada, M., Nishizawa, T., *et al.* (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.

Overbeek, R., Begley, T., Butler, R.M., Choudhuri, J.V., Chuang, H.Y., Cohoon, M., *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res* **33:** 5691–5702.

Papanikolaou, S., Galiotou-Panayotou, M., Fakas, S., Komaitis, M., and Aggelis, G. (2007) Lipid production by oleaginous Mucorales cultivated on renewable carbon sources. *Eur J Lipid Sci Technol* **109:** 1060–1070.

Partida-Martinez, L.P., and Hertweck, C. (2005) Pathogenic fungus harbours endosymbiotic bacteria for toxin production. *Nature* **437:** 884–888.

Partida-Martinez, L.P., and Hertweck, C. (2007) A gene cluster encoding rhizoxin biosynthesis in *Burkholderia rhizoxinica*, the bacterial endosymbiont of the fungus Rhizopus microsporus. *Chembiochem* **8:** 41–45.

Ploskon, E., Arthur, C.J., Evans, S.E., Williams, C., Crosby, J., Simpson, T.J., and Crump, M.P. (2008) A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. *J Biol Chem* **283:** 518–528.

Rambaut, A., M., Suchard, D., Xie, and A., Drummond. (2014). Tracer v1. 6. Available from http://beast.bio.ed.ac.uk/Tracer

Reysenbach, A.L., Giver, L.J., Wickham, G.S., and Pace, N.R. (1992) Differential amplification of rRNA genes by polymerase chain reaction. *Appl Environ Microbiol* **58:** 3417–3418.

Ruiz-Herrera, J., León-Ramírez, C., Vera-Nuñez, A., Sánchez-Arreguín, A., Ruiz-Medrano, R., Salgado-Lugo, H., *et al.* (2015) A novel intracellular nitrogen-fixing symbiosis made by Ustilago maydis and Bacillus spp. *New Phytol* **207:** 769–777.

Salvioli, A., Chiapello, M., Fontaine, J., Hadj-Sahraoui, A.L., Grandmougin-Ferjani, A., Lanfranco, L., and Bonfante, P. (2010) Endobacteria affect the metabolic profile of their host Gigaspora margarita, an arbuscular mycorrhizal fungus. *Environ Microbiol* **12:** 2083–2095.

Salvioli, A., Ghignone, Novero, S., Navazio, M., Bagnaresi, L.P., and Bonfante, P. (2015) Symbiosis with an endobacterium increases the fitness of a mycorrhizal fungus, raising its bioenergetic potential. *ISME J.* **10:** 130–144.

Sato, Y., Narisawa, K., Tsuruta, K., Umezu, M., Nishizawa, T., Tanaka, K., *et al.* (2010) Detection of Betaproteobacteria inside the mycelium of the fungus Mortierella elongata. *Microbes Environ* **25:** 321–324.

Schweizer, E., and Hofmann, J. (2004) Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. *Microbiol Mol Biol Rev* **68:** 501–517.

Sharma, M.S.M. (2008). A functional study on the multilateral symbiosis of the fungal order Sebacinales with plant hosts and bacteria. Ph.D. Dissertation. Justus-Liebig-Universität Gießen.

Simon, L., Bousquet, J., Lévesque, R.C., and Lalonde, M. (1993) Origin and diversification of endomycorrhizal fungi and coincidence with vascular land plants. *Nature* **363:** 67–69.

Soanes, D., and Richards, T.A. (2014) Horizontal gene transfer in eukaryotic plant pathogens. *Annu Rev Phytopathol* **52:** 583–614.

Spatafora, J.W., and Bushley, K.E. (2015) Phylogenomics and evolution of secondary metabolism in plant-associated fungi. *Curr Opin Plant Biol* **26:** 37–44.

Spatafora, J.W., Chang, Y., Benny, G.L., Lazarus, K., Smith, M.E., Berbee, M.L., *et al.* (2016) A phylum-level phylogenetic classification of zygomycete fungi based on genome-scale data. *Mycologia* **108:** 1028–1046.

Stajich, J.E., Berbee, M.L., Blackwell, M., Hibbett, D.S., James, T.Y., Spatafora, J.W., and Taylor, J.W. (2009) The fungi. *Curr Biol* **19:** R840–R845.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30:** 1312–1313.

Stevens, M.P., Haque, A., Atkins, T., Hill, J., Wood, M.W., Easton, A., *et al.* (2004) Attenuated virulence and protective efficacy of a Burkholderia pseudomallei bsa type III secretion mutant in murine models of melioidosis. *Microbiology* **150:** 2669–2676.

Talavera, G., and Castresana, J. (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* **56:** 564–577.

Taylor, T.H., Krings, M., and Taylor, E.D. (2015) Fossil fungi. *Nord J Botany* **33:** 126–127.

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N.S., Wijesundera, R., *et al.* (2014) Global diversity and geography of soil fungi. *Science* **346:** 256688.

Tisserant, E., Malbreil, M., Kuo, A., Kohler, A., Symeonidi, A., Balestrini, R., *et al.* (2013) Genome of an arbuscular mycorrhizal fungus provides insight into the oldest plant symbiosis. *Proc Natl Acad Sci U S A* **110:** 20117–20122.

Torres-Cortes, G., Ghignone, S., Bonfante, P., and Schussler, A. (2015) Mosaic genome of endobacteria in arbuscular mycorrhizal fungi: Transkingdom gene transfer in an ancient mycoplasma-fungus association. *Proc Natl Acad Sci U S A* **112:** 7785–7790.

Tschaplinski, T.J., Standaert, R.F., Engle, N.L., Martin, M.Z., Sangha, A.K., Parks, J.M., *et al.* (2012) Down-regulation of the caffeic acid O-methyltransferase gene in switchgrass reveals a novel monolignol analog. *Biotechnol Biofuels* **5:** 71.

Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley, R.E., *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature* **457:** 480–484.

Vannini, C., Carpentieri, Salvioli, A., Novero, A., Marsoni, M., Testa, M.L., *et al.* (2016) An interdomain network: the endobacterium of a mycorrhizal fungus promotes antioxidative responses in both fungal and plant hosts. *New Phytol.* **211:** 265–275.

Wang, L., Chen, W., Feng, Y., Ren, Y., Gu, Z., Chen, H., *et al.* (2011) Genome characterization of the oleaginous fungus *Mortierella alpina*. *PLoS One* **6:** e28319.

Wattam, A.R., Abraham, D., Dalay, D.O., Disz, T.L., Driscoll, T., Gabbard, J.L., *et al.* (2014) PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Res* **42:** D581–91.

Yao, J., and Rock, C.O. (2013) Phosphatidic acid synthesis in bacteria. *Biochim Biophys Acta* **1831:** 495–502.

Yao, Z., Davis, R.M., Kishony, R., Kahne, D., and Ruiz, N. (2012) Regulation of cell size in response to nutrient availability by fatty acid biosynthesis in *Escherichia coli*. *Proc Natl Acad Sci* **109:** E2561–E2568.

Zhang, Y.M., and Rock, C.O. (2016). Chapter 3 - Fatty acid and phospholipid biosynthesis in prokaryotes. In *Biochemistry of*

Lipids, Lipoproteins and Membranes. McLeod, N.D.R.S. (ed) (Sixth Edition). Boston: Elsevier, pp. 73–112.
Zhou, Y.J., Buijs, N.A., Siewers, V., and Nielsen, J. (2014) Fatty acid-derived biofuels and chemicals production in Saccharomyces cerevisiae. *Front Bioeng Biotechnol* **2:** 1–6.

Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Table S1. Fungal isolates screened for bacteria through 16S rDNA amplification and sequencing
Table S2. Genome size, assembly and annotation statistics for *Mortierella elongata* (AG77) and its associated endobacterium *Mycoavidus cysteinexigens* (AG77).
Table S3. A. CAZys detected in the genome of Mortierella elongata (AG77). B. CAFE analysis results of the top 40 protein families in expansion (excluding transposon-related families) in *Mortierella elongata* genome as compared to representative fungi. Aspni: *Aspergillus niger*; Neucr: *Neurospora crassa*; Ustma: *Ustilago maydis*; Copci: *Coprinopsis cinerea*; Rhiir: *Rhizophagus irregularis*; Catan: *Catenaria anguillulae*; Batde: *Batradochytridium dendrobaditis*; Phybl: Phycomyces blakesleeanus; Rhior: *Rhizopus oryzae*; Morve: *Mortierella verticillata*; Morel : *Mortierella elongata*; Umbra: *Umbelopsis ramanniana*.
Table S4. A. Genes involved in *Mortierella elongata* involved in fatty acid and glycerolipid synthesis. Homologous proteins were identified as having >65% amino acid (aa) sequence identity and <100 aa difference in sequence length. Italics indicate potential homologs at 50-65% sequence identity or >65% sequence identity and 100-200 aa difference in sequence length. Percent identity is given in parentheses next to each homolog.
Table S5. Comparative analysis of parallel gene loss between fungal endosymbionts in *Burkholderia, Ca*. Glomeribacter, and *Mycoavidus*.
Table S6. Type 3 Secretion System (T3SS) components from *Mycoavidus cysteinexigens* genome.
Table S7. Putative non-ribosomal peptide synthase (NRPS) gene clusters identified in *Mycoavidus cysteinexigens*. File is GenBank formated.
Table S8. Specific primers for *Mycoavidus cysteinexigens* 16S rDNA reported 5′ to 3′.
Table S9. Metabolomics of strains of *Mortierella elongata* with endosymbionts and cleared of endosymbionts grown with and without peptone.
Table S10. Volatile compounds were differentially produced among *Mortierella elongata* AG77 with and without bacteria). Identified compounds included alcohols, ketones, an aldehyde, a sulfur containing volatile, a furane as well as a series of unidentified volatiles assigned to sesquiterpenoids given their characteristic fragments (i.e. m/z 204). The presence of the bacterial symbiont induced differences compared to the strain without the bacteria. Volatiles which differed statistically ($P<0.05$) between "+ bact" and "−bact" strains are reported in the right column.
Fig. S1. A. Genome scale RAxML phylogeny illustrating relationships between Mortierella elongata and other fungi with sequenced genomes. Numbers above nodes indicate bootstrap support for taxon groupings. Asterisks (*) indicate bootstrap support of 100%. Branch colors indicate and black bars on the far right indicate phylum level designations. **B.** KEGG/pfam analysis tree**. C.** Functional comparison of the PFAM protein domains of *M. elongata* with twelve other fungi. The top 100 PFAM domains found in all genomes were selected. The frequency values were transformed into z-scores, which are measure of relative enrichment (red) and depletion (green); the hierarchical clustering was done with a Euclidian distance metric and average linkage clustering method. Taxa are named as followed: Morve1 = *Mortierella verticillioides*; Morel2=*Mortierella elongata*; Umbra1 = *Umbelopsis rammaniana*; Phybl2 = *Phycomyces blakesleeanus*; Rhior3 = *Rhizopus oryzae*; Pucgr2 = *Puccinia graminis*; Ustmal = *Ustilago maydis*; Neucr1 = *Neurospora crassa*; Copci1 = *Coprinopsis cinerea*; Catan1 = *Catenaria anguillulae*; Batde5 = *Batrachochytrium dendrobatidis*; Rhiir2 = *Rhizophagus irregularis*; Aspni7 = *Aspergillus nidulans*.
Fig. S2. Biolog plates confirm unique substrate utilization patterns: 1 N-acetyl glucosamine (chitin degradation – 261 genes), 2 Tween-80 (lipid degradation, positive – 76 lipase genes), 3 Cellobiose (cellulose degradation, negative)
Fig. S3. Genome of circular chromosome of Mortierella elongata endosymbiont, *Mycoavidus cysteinexigens* (AG77). From the outer ring in tracks represent physical location, CDs leading and then lagging, rRNAs and tRNAs, and GC content skew
Fig. S4. Dotplot and alignment
Fig. S5. The pentose-phosphate shunt is completely missing in both *Ca*. Glomeribacter gigasporarum and the Glomeribacter-related endosymbiont (*Mycoavidus cysteinexigens* AG77) of *Mortierella elongata*
Fig. S6 AntiSMASH analysis secondary metabolite predicted gene cluster 1: predicted sideropohore locus between 551795-563831bp. Query sequence from *Mycoavidus cysteinexigens* genome not shown, taxa with similar gene clusters shown in descending order of homology below, arrows indicate gene orientation, colors indicate homologous genes.
Fig. S7. AntiSmash predicted secondary metabolite gene cluster 2: predicted monomer of axinastatin at locus 722832-76142. Query sequence from *Mycoavidus cysteinexigens* genome not shown, taxa with similar gene clusters shown in descending order of homology below, arrows indicate gene orientation, colors indicate homologous genes.
Fig. S8. AntiSmash predicted secondary metabolite gene cluster 3: predicted non-ribosomal peptide synthetase at locus 1129074-1178025. Query sequence from *Mycoavidus cysteinexigens* genome not shown, taxa with similar gene clusters shown in descending order of homology below, arrows indicate gene orientation, colors indicate homologous genes.
Fig. S9. AntiSmash predicted secondary metabolite gene cluster 4: predicted as capable of producing arylopolyene at locus 2217060-2258265. Query sequence from *Mycoavidus cysteinexigens* genome not shown, taxa with similar gene clusters shown in descending order of homology below, arrows indicate gene orientation, colors indicate homologous genes.
Fig. S10. AntiSmash predicted secondary metabolite gene cluster 5: predicted as contributing to synthesis of lassopeptide at locus 2312505-2334986. Query sequence from

Mycoavidus cysteinexigens genome not shown, taxa with similar gene clusters shown in descending order of homology below, arrows indicate gene orientation, colors indicate homologous genes.

Fig. S11. AntiSmash predicted secondary metabolite gene cluster cluster 6: predicted non-ribosomal peptide synthetase gene cluster at locus 2393151-2438943. Query sequence from *Mycoavidus cysteinexigens* genome not shown, taxa with similar gene clusters shown in descending order of homology below, arrows indicate gene orientation, colors indicate homologous genes.

Fig. S12. 16S phylogeny of Burkholderiaceae. Outgroup taxa are shown in black. *Burkholderia* is shown in purple. *Ca*. Glomeribacter seqeuenced from fungi within the Gigasporaceae and are shown in green. *Mycoavidus* spp. are shown in blue. The genome isolate (AG77) is labelled in red.

Fig. S13. Fingerprints of significant VOC emission factors of cleared (red) and uncleared strain (pink) AG77 and media controls (green) over time. VOC emission factors for significant products are identified below each graph.

Fig. S14. Fingerprints of CO2 VOC emission factors of cleared (red) and uncleared strain (pink) AG77 and media controls (green). *Mortierella elongata* AG77 respired approximately by a factor of 2 more CO2. There were up to 30% differences in "Plus" replicates, but the respirations in "Minus" were almost identical. Media did not respire but may have absorbed small amount of CO2.

Fig. S15. A. Primer map for *Mycoavidus cysteinexigens* specific 16S rDNA primers. B Gel image of PCR products of *M. cysteinexigens* specific primer combinations. Primer combinations are listed across the top. A 1kb ladder was used on the far left. the bands in the left panel are from M. elongata strain AG77 with *M. cysteinexigens*, the panel on the right are PCR reactions from the *Mortierella elongata* AG77 strain exposed to antibiotics

Fig. S16. Schematic of the experimental VOC emission measurement setup. All outgoing sampling lines from the incubator and the stream selection valve were heated to 60 $^{\circ}$C except for the subsampling line to instruments measuring CO_2 and relative humidity. The multiport stream selection flow through valve cycled between sampling the 8 individual jars. Humidity was controlled in the supply air to the incubated jars from the Zero Air Generator (ZAG) by mixing flows from the water saturated and dry lines.