The complexities of inferring symbiont function: *Paraburkholderia* symbiont dynamics in social amoeba populations and its impact on the amoeba microbiome

Authors

James G. DuBose[1], Michael S. Robeson[2], Mackenzie Hoogshagen[1], Hunter Olsen[1], Tamara S. Haselkorn[1*]

Affiliations

[1]Department of Biology, University of Central Arkansas, 201 Donaghey Avenue, Conway, AR 72035

[2]Department of Biomedical Informatics, University of Arkansas for Medical Sciences, Little Rock, AR, 72205

*corresponding author email: thaselkorn@uca.edu

The authors declare no competing financial interests

Abstract

The relationship between the social amoeba *Dictyostelium discoideum* and its endosymbiotic bacteria *Paraburkholderia* provides a model system for studying the development of symbiotic relationships. Laboratory experiments have shown that any of three species of *Paraburkholderia* symbiont allow *D. discoideum* food bacteria to persist through the amoeba lifecycle and survive in amoeba spores, rather than being fully digested. This phenomenon is termed “farming”, as it potentially allows spores dispersed to food poor locations to grow their own. The occurrence and impact of farming in natural populations, however, has been a challenge to measure. Here, we surveyed natural *D. discoideum* populations and found that only one of the three symbiont species, *P. agricolaris,* remained prevalent. We then explored the effect of *Paraburkholdia* on the amoeba microbiome, expecting that by facilitating bacterial food carriage it would diversify the microbiome. Contrary to our expectations, *Paraburkholderia* tended to infectiously dominate the *D. discoideum* microbiome, in some cases decreasing diversity. Similarly, we found little evidence for *Paraburkholderia* facilitating the carriage of particular food bacteria. These findings change our understanding of farming and suggest the possibility that *Paraburkholderia* could be playing multiple roles for its host, as inferred metagenomic analysis indicates a potential role of *P. agricolaris* in toxin degradation.

Introduction

Bacterial symbionts have been found to associate with hosts across the tree of life, with often dramatic effects on their ecology and evolution. The roles of bacterial symbionts within their hosts can range from mutualism to parasitism and can vary based on host and environmental contexts, as well as on the length of the association and mode of transmission (1–3). Amoebae, a polyphyletic group of unicellular eukaryotes, serve as promising model systems for studying symbiotic relationships, particularly the bacterial transition from free-living to intracellular. When amoebae feed on bacteria, a selective pressure is exerted that favors bacteria with mechanisms to resist digestion (4–6). Hence, amoebae serve as intracellular training grounds for bacteria to evolve phagocytosis resisting mechanisms, allowing certain bacteria to occupy amoebae as endosymbionts (7–9). Bacterial endosymbionts have been commonly identified in *Acanthamoeba* hosts, including bacterial lineages of Chlamydiae, *Proteobacteria*, *Bacteroidetes*, and *Dependentiae* (10–12), although their effects on their hosts are often unclear. For example, while Chlamydiae symbionts are capable of living as energy parasites, suggesting symbionts are costly to their hosts (13), other Chlamydiae species protect *Acanthamoeba* from the pathogenic *Legionella* bacteria and enhance cell motility and growth, suggesting symbionts are beneficial to their hosts (14–17). The difficulty of studying the microbial community dynamics in natural environments dampens our understanding of the ecological contexts of amoeba – bacteria interactions, which makes inferring the function of a particular symbiont in its host a challenge.

The social amoeba *Dictyostelium discoideum* has gained traction as a model organism to study host cell – bacteria interactions in laboratory settings. Social amoebae are soil-dwelling protists that have unique life cycles that involves the transition from a single-cellular state to a multicellular state when their bacterial food source is depleted (18). Tens of thousands amoeba cells aggregate to ultimately form a fruiting body, which consists of two primary structures: the sorus, a structure used to house spores, and the stalk, which is composed of dead cells that differentiated to support the sorus (19). It was previously thought that sentinel cells, which perform immune-like functions, cleared all bacteria from the amoeba fruiting body throughout the social cycle (20), as evidenced by the inability to culture bacteria from many laboratory-isolated *D. discoideum* clones. Recent studies, however, have shown that in nature *D. discoideum* harbors a microbiome consisting of both transient and persistent bacterial symbionts (21,22), belonging to the phyla Proteobacteria, Actinobacteria, Bacteroidetes, Firmicutes, Chlamydiae, and Acidiobacteria. This recent characterization of the *D. discoideum* microbiome has expanded its potential to serve as a model system for studies pertaining to microbiome formation and maintenance (23) and some evidence suggests that *D. discoideum* may play an active role in its microbiome formation by secreting lectin proteins that bind bacteria during the social cycle (24).

In natural populations of *D. discoideum*, the most widespread and persistent bacterial symbionts are *Paraburkholderia*, Chlamydiae, and *Amoebophilus* (25). Although the functions of these symbionts in natural populations are largely unknown, the

Paraburkholderia farming symbionts have been the best characterized. The phenomenon of farming was demonstrated primarily in laboratory experiments by comparing the ability of *Paraburkolderia* infected and uninfected *D. discoideum* to grow when a sorus was plated without food bacteria. *Paraburkholderia* are inedible, but they allow bacterial food carriage that is presumed to benefit the amoeba in nature if the spores are dispersed to an area that is scarce in food bacteria, as the soil could be seeded with the carried bacteria. There is, however, a cost to infection, as *Parburkholderia*-infected *D. discoideum* produce fewer spores in food rich laboratory conditions (26,27). While many members of the *Paraburkholderia* and closely related *Burkholderia* genera form symbiotic interactions with a variety of plant, fungi, and insect hosts (28–30), only three *Paraburkholderia* species, *P. agricolaris*, *P. bonniea*, and *P. hayleyella* have been found to confer farming and form persistent symbiotic relationships with *D. discoideum* (31). The variability in fitness costs and the extent of food carriage that each symbiont permits makes this *Paraburkholderia – D. discoideum* symbiosis context dependent (32), like many symbiotic relationships (33). Screenings for *Paraburkholderia* in natural *D. discoideum* populations have shown that *P. agricolaris* tends to be the most prevalent, followed by *P. hayleyella* and *P. bonniea* (31*)*. Indeed, these differences in symbiotic potential likely translate to different host outcomes, which could be responsible for the differing levels of prevalence between each *Paraburkholderia* species in nature.

Measurements of fitness in laboratory settings have aided our understanding of the consequences of *Paraburkholderia* infection, but its role in natural populations is less

clear. To better understand the occurrence and impact of this of this symbiosis, we used PCR based screenings of natural *D. discoideum* populations to identify current symbiont prevalence, transmission patterns, and genetic diversity. Since symbiont transmission modes play a large role in determining whether a symbiont's evolutionary trajectory will lead to mutualism, benign parasitism, or extreme pathogenicity (3), we looked for evidence showing whether *P. agricolaris* was environmentally acquired or vertically transmitted between social amoeba hosts. To determine if *Paraburkholderia* was conferring bacterial food carriage in nature, we compared the microbiomes of infected and uninfected *D. discoideum*. We hypothesized that if *Paraburkholderia* was indiscriminately facilitating the carriage of food bacteria, infected amoeba would possess more diverse microbiomes. Alternatively, if *Paraburkholderia* was facilitating the carriage of certain food bacteria, perhaps indicative of *D. discoideum* feeding preferences, we hypothesized that *Paraburkholderia* infected amoeba would have microbiomes that were similar in composition to each other. Overall, our results are inconsistent with the idea that *Paraburkholderia* is a co-evolved mutualist that facilitates bacterial food carriage and suggest an alternative relevance to *D. discoideum* in nature.

Methods

Soil Collection and Amoeba Culturing

We re-sampled natural *D. discoideum* populations in the Virginia Mountain Lake area (25,31). Soil, sampled from below the leaf litter layer of the forest floor or below rotting logs, was collected from ten different sites at each of four separate locations

(Supplemental Table 1) and was plated on hay agar within 96 hours (Supplemental Table 2). After 4-5 days of growth time, Dictyostelid species were morphologically identified via stereomicroscopy (34), and 5-10 sori from each patch were collected with sterile pipette tips and suspended into 250µl of KK2 spore buffer. Every patch of *D. discoideum*, identified by the presence of a basal disc at the base of its stalk, was collected, as well as other social amoeba species, including *D. mucoroides*, *D. giganteum*, *D. purpureum*, *Heterostelium pallidum*, and *Polysphondylium violaceum*. We then resuspended fruiting bodies in 100µl of KK2 buffer for "passaging": replating onto SM/5 agar (Supplemental Table 2) with the addition of 200µl of a 1.5 OD_{600nm} suspension of *Klebsiella pneumoniae* food bacteria to increase the quantity of fruiting bodies. Subsets from each amoeba species were sequenced to confirm amoeba identity (See following section).

<u>Symbiont Screening</u>

DNA from 15-20 fruiting bodies from the passaged SM/5 plates was extracted using the a Chelex/proteinaseK protocol as in (31). PCRs were conducted to screen each sample for *Dictyostelium* DNA by amplifying approximately 500 base pairs of the 18S rRNA gene using the D307F forward primer and the D862R reverse primer (35). To screen for *Paraburkholderia*, we used the *Paraburkholderia* specific primers Burk3F and Burk 3R, which amplify a small section of the 16S rRNA gene (36), and positive samples were sent for Sanger sequencing at Eurofins Genomics. Sequences were cleaned using Geneious v8 (http://www.geneious.com) (37), and NCBI Blast was used to classify *Paraburkholderia* species. In samples that were identified as either *P. agricolaris*, *P.*

bonniea, or *P. hayleyella*, the LepA gene, which allows for better species level identification and haplotype resolution (31), was then amplified and sequenced using the LepA forward and reverse primers (38). We screened for Chlamydiae using forward primer CHL16SFOR2 and the reverse primer CHL16SREV2 (39), and for *Amoebophilus* using the forward primer AmoF and reverse primer AmoR (25)(Supplemental Table 3).

P. agricolaris phylogeny, Population Genetics, and Co-infection Statistics

LepA sequences were used to construct a phylogenetic tree of *P. agricolaris* haplotypes. Overlapping peaks in the DNA chromatogram at a single site were indicative of co-infections with multiple haplotypes and were excluded from analysis because haplotype designations were not able to be resolved. A Maximum Likelihood phylogenetic tree was constructed in MEGA7 using the Tamura 3-parameter+G+I model with 1 000 bootstrap replicates (40). Previously identified LepA haplotypes were also placed to determine if these haplotypes had previously been previously identified (31).

To examine *P. agricolaris* genetic differentiation between *D. discoideum* populations, pairwise F_{ST} values were calculated in DNAsp (v6) (41) using non-coinfected LepA sequences (n=65). Concatenated 16S rRNA and LepA sequences were then used to look for evidence of recombination using the Recombination Detection Program (RDP) (v4) (42). Pairwise Fisher exact tests (2x2 contingency tables) between *P. agricolaris*, Chlamydiae, and *Amoebophilus*, were used to investigate ecological co-infection tendencies between symbionts.

Laboratory inoculations of *D. discoideum* with different *Paraburkholderia* species

Separate *P. agricolaris, P. bonniea*, and *P. hayleyella* solutions were made by resuspending stationary phase bacterial colonies from SM/5 plates into KK2 buffer at a concentration of 1.5 OD_{600nm} and combined with 1.5 OD_{600nm} *K. pneumoniae* suspensions at a final volume of 95%/5% *K.pneumoniae*/*Paraburkholderia.* We then plated 10^5 *D. discoideum* spores (strain QS864) with 200µl of each *Paraburkholderia* solution onto SM/5 plates and incubated at room temperature. Amoebae were collected after one week of growth time, and *Paraburkholderia* was cultured from fruiting bodies to confirm infection. Ten fruiting bodies from each plate were resuspended in KK2 spore buffer and passaged again to confirm persistent *Paraburkholderia* infection. *Paraburkholderia* infected *D. discoideum* lines were then innoculated into 2.5 grams local soil collected from a natural area maintained by the Department of Biology at the University of Central Arkansas (Conway, AR) overnight. Soil contents were subsequently plated as describe above, and *D. discoideum* fruiting bodies were collected directly from these soil pates after one week of growth time.

DNA Extractions for Microbiome Analyses

Approximately 50-100 sori were collected from passaged SM/5 plates (natural infections) or soil plates (lab infections), placed into Hyclone water, and stored at -80°C until DNA was extracted. DNA extractions were performed using the Qiagen DNeasy PowerSoil Pro Kit and the Qiagen DNeasy UltraClean Pro Kit, according to

manufacturer's protocols. Samples were then sent for sequencing to Argonne National Laboratories (Lemont, IL, USA). Here, samples were sequenced on the Illumina MiSeq platform using 250bp paired-end cycles. Library preparation and sequencing was performed using the 799F-1115R primer set, which amplifies the V5-V6 region of the 16S rRNA gene (43), as the more commonly used primers for the V4 region preferentially amplify amoeba mitochondrial DNA (22). Sequences were returned as FASTQ files with adaptors removed.

Illumina Sequence Processing

Illumina sequence files were demultiplexed and analyzed in QIIME2 (v2020.6). Chimeric sequences were removed, primers were trimmed, and paired-ends were joined using q2-dada2 plugin [45–48] (44–47). The q2-RESCRIPt plugin was used to curate the SSU SILVA NR99 (version 138) database, and to train a custom V5-V6 amplicon specific classifier (48). The classify-sklearn method was used in the q2-feature-classifier plugin to assign taxonomic names to amplicon sequence variants (ASVs) (49–51). ASVs classified as either mitochondria or chloroplast, as well as ASVs that did not classify to the phylum-level were removed for downstream analysis. ASVs with a percent identity/alignment of less than 0.80/0.80 to the SILVA reference sequences were excluded using the q2-quality-control plugin (52). A phylogenetic tree for downstream analyses was created by inserting ASVs into the Silva 128 reference database using SATé-enabled phylogenetic placement (SEPP), where all fragments were successfully inserted (50). Although the V5-V6 region of 16S rRNA did not provide species level resolution, we identified the genus *Burkholderia-Caballeronia-Paraburkholderia* to

represent *P. agricolaris* based on Sanger sequences of the more variable LepA gene in our naturally infected study. In laboratory inoculated samples, the genus *Burkholderia-Caballeronia-Paraburkholderia* was identified to represent both *P. agricolaris* and *P. hayleyella* and the genus *Pandorea* was identified to represent *P. bonniea*, based on their high abundance in their respective infection groups.

Microbiome Diversity Analyses

After constructing alpha-rarefaction curves, natural samples were rarefied to a depth of 3 950 and laboratory inoculated samples were rarefied to a depth of 4 000 (53). This allowed us to retain 28 *P. agricolaris* infected 16 uninfected natural samples and 19 *P. agricolaris*, 9 *P. bonniea*, and 10 *P. hayleyella* artificially infected samples. Alpha-diversity was evaluated by using the q2-diversity plugin to calculate average ASVs, Pielou's evenness, and Faith's phylogenetic diversity (50,54). Significance of each metric was evaluated using Kruskal-Wallis and pairwise Kruskal-Wallis tests (α=0.05 with FDR q-values) (55). Beta-diversity was evaluated by using the q2-deicode plugin to calculate the Aitchison distances between non-rarefied sample groups (56). Weighted UniFrac and Unweighted UniFrac were also used on rarefied data to evaluate beta-diversity using the q2-diversity plugin (57–61). Ordination plots were generated for each beta-diversity metrics and between group significance was evaluated using PERMANOVA and PERMDISP with 999 permutations and FDR adjusted p-values (62). When PERMAONVA and PERMDISP results were both significant, ordination plots were consulted to determine separation between groups. The feature table generated

by QIIME2 was uploaded to microbiomeanalyst.ca to perform co-association analyses using the SparCC method.

Microbiome Functional Inference

The Phylogenetic Investigation of Communities by Reconstruction of Unobserved States 2 (PICRUSt2) pipeline (63) was used in QIIME2 (v2019.10) to estimate the functional potential of the natural *D. discoideum* microbiome. The feature table and representative sequences from diversity analyses were used to predict functional profiles via gene family abundance. Pathway abundances were inferred from the predicted KEGG Ortholog (KO) abundances. Univariate comparisons were conducted to search for enrichments or depletions in different metabolic pathways between sample groups. Pathway abundances were analyzed using microbiomeanalyst.ca to perform Mann-Whitney/KW test, with FDR adjusted p-values.

Results

P. agricolaris remains prevalent in natural *D. discoideum* populations but does not co-associate with Chlamydiae or *Amoebophilus* endosymbionts

To identify the current prevalence of *D. discoideum* endosymbionts, we screened 244 *D. discoideum* isolates collected from the Mountain Lake area in 2019 for *Paraburkholderia,* Chlamydiae, and *Amoebophilus*. We then compared the prevalence of each endosymbiont in 2019 to that of a previous collection in 2000 (25,31). *P.*

agricolaris exhibited a prevalence of 35.1%, up from a 14.9% prevalence in 2000. *P. bonniea* was only found once in our sampling, resulting in a prevalence of 0.79% (down from 2.6% in the 2000 collection). *P. hayleyella*, which was previously identified to be 8% prevalent in 2000, was not found in our collection. *P. agricolaris* appeared to be stochastically distributed between each location and collection site, where prevalence varied from less than 15% at location 1 to greater than 60% at location 4 (Figure 1, Supplemental Table 2). Chlamydiae, which was 19% prevalent in 2000, was found to be 58% prevalent in our collection. *Amoebophilus*, 37% prevalent in 2000, was found to be 18% prevalent in our collection. Pairwise Fishers exact tests were used to look for signs of positive or negative co-associations between *P. agricolaris*, Chlamydiae, and *Amoebophilus* at each location. A positive association was observed between *P. agricolaris* and Chlamydiae at location 3 (p=0.0472, 23 expected with 26 observed); however, no other significant associations were found (Supplemental Table 4).

P. agricolaris is genetically diverse and shows evidence of recombination and environmental acquisition

To gain insight into the natural and co-evolutionary history of *P. agricolaris* and *D. discoideum,* we investigated *P. agricolaris* haplotype diversity, genetic recombination tendencies, and transmission patterns. Phylogenetic analyses of LepA sequences (n=65) revealed ten distinct haplotypes, and placement of previous *P. agricolaris* sequences showed that 7 of these haplotypes have not been previously identified (Figure 2a) (31). Haplotype 7 was the only haplotype found in amoebae from multiple locations and pairwise F_{ST} comparisons showed that the locations harboring haplotype

7 were the most similar in terms of *P. agricolaris* genetic composition (Supplemental table 4). Furthermore, up to 5 different haplotypes were identified in amoebae from a single soil sample (Supplemental Table 5). Using concatenated 16S rRNA and LepA sequences we identified two putative inter-locus recombination events, both between haplotypes 2 and 9. We used the presence of a *P. agricolaris* haplotype infecting multiple amoebae species as evidence of environmental acquisition. In addition to infecting *D. discoideum,* haplotype 3 was identified in *D. purpureum,* haplotype 4 was identified in *D. mucoroides*, haplotype 8 was identified in *H. pallidum*, and haplotype 7 was identified in *D. giganteum, H. pallidum, P. violaceum*, and *D. mucoroides* (Figure 2).

Paraburkholderia symbionts dominate the *D. discoideum* microbiome and cause a general decrease in microbiome diversity

Our investigation of the impacts of *Paraburkholderia* on the natural *D. discoideum* microbiome was limited to *P. agricolaris*, as *P. bonniea* and *P. hayleyella* were not found. Overall, denoising of 16S rRNA sequences resulted in 130 amplicon sequence variants (ASVs) with a total frequency of 1 573 424 reads. These were represented by 6 phyla, 9 classes, 33 orders, 49 families, and 57 genera. The majority of sequences belonged to the genus *Burkholderia-Caballeronia-Paraburkholderia* (57.16%), which was identified to represent *P. agricolaris* based on Sanger sequences of the more variable LepA gene. Although we did not find *P. hayleyella* or *P. bonniea* infected *D. discoideum* in nature, we were able to gain insights on their impacts by culturing artificially infected *D. discoideum* on local soil. This yielded 168 ASVs with a total

frequency of 2 135 808 reads. These were represented by 10 phyla, 15 classes, 49 orders, 77 families, and 140 genera.

If *Paraburkholderia* infection was somehow facilitating the carriage of random environmental bacteria, we posited that *Paraburkholderia*-infected *D. discoideum* would have a more diverse microbiomes than uninfected ones. However, we found that *P. agricolaris* tended to dominate the *D. discoideum* microbiome, accounting for an average of 71.35% of bacterial community abundance, for greater than 90% of reads in the majority (16/28) of infected samples, and for 76.86% of all reads (Figure 3a). The *D. discoideum* microbiome evenness and richness was not significantly altered by *P. agricolaris* infection (Pielou's Evenness, $p=0.213399$; observed ASVs, $p=0.376987$). However, *D. discoideum* infected with *P. agricolaris* did possess less phylogenetically diverse microbiota than uninfected *D. discoideum* (Faith's Phylogenetic Diversity, $p=0.048117$) (Figure 4a). Laboratory inoculation of each *Paraburkholderia* species into *D. discoideum* yielded similar results. *P. agricolaris*, *P. bonniea*, and *P. hayleyella* each had abundances of greater than 90% in 13/19, 8/9, and 10/10 infected *D. discoideum* isolates, respectively (Figure 3b). Consistent with the trend in *Paraburkholderia* abundance, laboratory inoculation with each of the *Paraburkholderia* species resulted in significant reductions in *D. discoideum* microbiome evenness, richness, and phylogenetic diversity (Pielou's Evenness, $p \leq 0.001515$; observed features, $p \leq 0.041061$; Faith's Phylogenetic Diversity, $p \leq 0.041759$) (Figure 4b). *P. hayleyella* infection significantly reduced microbiome evenness when compared to *P. agricolaris* (p

= 0.041457), but no other pairwise differences were observed between each *Paraburkholderia* species (Figure 4b).

Overabundance of *Paraburkholderia* alters the composition and structure of the *D. discoideum* microbiome

While *Paraburkholderia* infection did not diversify the *D. discoideum* microbiome, we thought it possible that *Paraburkholderia* might facilitate the carriage of specific food bacteria, altering the microbiome structure and/or composition. ß-diversity comparisons using weighted UniFrac, which measures microbiome structure by considering the phylogenetic relatedness and relative abundances of each ASV, showed that the microbiomes of infected *D. discoideum* were similar to each other, and dissimilar to the microbiomes of uninfected *D. discoideum* (PERMANOVA, pseudo-F=10.5405, p=0.001; PERMDISP, F=8.97892, p=0.01; where significant PERMDISP results were negated by the clear grouping of *P. agricolaris* infected microbiomes in PCoA space) (Figure 5a). This is not surprising given the high relative abundance of *Parburkholderia*. ß-diversity comparisons using unweighted UniFrac, which measures microbiome composition by only considering phylogenetic relatedness and presence or absence of ASVs showed the effect of *Paraburkholderia* without directly considering its relative abundance. Here we also saw significant differences between *P. agricolaris* and infected and uninfected amoebae (PERMANOVA, pseudo-F=3.3698, p=0.002; PERMDISP, F=14.0089, P=0.003 where significant PERMDISP results were negated by a clear distinction between infected and uninfected samples upon PCoA visualization). (Figure 5b). The impact, however, is unlikely due to carriage of specific food bacteria, but rather that the

high relative abundance of *P. agricolaris* corresponds with a reduction in microbiome phylogenetic diversity (Figure 4a).

The contribution of *P. agricolaris* to the similarity between infected *D. discoideum* samples is more quantitatively highlighted via Robust Aitchison analysis, which is a novel beta-diversity metric that is more adept at handling the sparsity of microbiome data. The compositional nature of Robust Aitchison analysis also makes it proficient at revealing salient members of microbial communities that are responsible for sample clustering [60]. Along with showing significant differences between *P. agricolaris* infected and uninfected *D. discoideum* microbiomes (PERMANOVA, pseudo-F=9.28284, p=0.001; PERMDISP, F=0.655692, p=0.392) Robust Aitcheson PCA highlights the influence of *P. agricolaris* abundance (indicated by labeled vector) on the grouping of infected *D. discoideum* microbiomes (Figure 5c). Because this metric takes into account actual abundances rather than relative abundances, it suggests that *Paraburkholderia* is outcompeting other bacteria rather than simply adding to the overall size of the microbiome.

D. discoideum inoculated with *Paraburkholderia* in the laboratory showed similar, yet more pronounced, differences when compared to uninfected *D. discoideum*. *D. discoideum* microbiome structure, determined using weighted UniFrac, was significantly altered by infection with *P. agricolaris*, *P. hayleyella*, or *P. bonniea* when compared to uninfected *D. discoideum* (PERMANOVA, pseudo-F ≥ 7.384689, p ≤ 0.007,

PERMDISP, F=26.8176, p=0.001 where significant PERMDISP results were negated by clear PCoA grouping) (Figure 6a). Likewise, *D. discoideum* microbiome composition, evaluated using unweighted UniFrac, was altered by infection with either *P. agricolaris*, *P. hayleyella*, or *P. bonniea* (PERMANOVA, pseudo-F ≥ 4.200470, p ≤ 0.002, PERMDISP, F=2.30245, p=0.11) (Figure 6b). Microbiome compositional comparisons using Aitchison distances showed a clear distinction between the microbiomes of *Paraburkholderia* infected amoebae and uninfected amoebae (PERMANOVA, pseudo-F=22.282, p=0.001, PERMDISP, F=8.9081, p=0.001). Furthermore, Aitchison PCA highlights that the abundance of each *Paraburkholderia* species is responsible for their impact on microbiome composition (Figure 6c). These results are consistent with the reduction in evenness, richness, and phylogenetic diversity observed in α-diversity comparisons.

P. agricolaris positively associates with *Rhizobiales* in the natural *D. discoideum* microbiome

Although we did not find indication of an enhancement in bacterial food carriage by *Paraburkholderia* via compositional comparisons, we reasoned that food carriage could be indicated by specific co-associations with *P. agricolaris.* Therefore, we performed co-association analyses using the SparCC method, which accounts for bacterial presence and abundance, and is adept at handling the sparsity of microbiome data. We identified only one positive association between *P. agricolaris,* represented by the order Burkholderiales, and the order Rhizobiales (p=0.0297), and no significant negative

association. Overall, significant associations between other members of the amoeba microbiome were scarce (Figure 7).

Chlamydiae is abundant in the natural *D. discoideum* microbiome but does not affect its diversity

Given the high prevalence of Chlamydiae in *D. discoideum* populations, we investigated how it might impact the natural *D. discoideum* microbiome. We found that the phylum Chlamydiae accounted for 21.57% of all sequence reads, making it the second most abundant group. The average abundance of Chlamydiae across infected samples was 43.34%, where Chlamydiae accounted for 36.86% of all reads in infected samples and for greater than 90% of reads in 8/25 infected samples (Figure 3a). Chlamydiae infection was not correlated with microbiome evenness, richness, or phylogenetic diversity (Pielou's Evenness, p=0.731175; observed features, p=0.257608; Faith's Phylogenetic Diversity, p=0.213492). However, Chlamydiae infected *D. discoideum* did have distinct microbiome compositions and structures, measured by Weighted UniFrac (PERMANOVA, pseudo-F=5.38621, p=0.004, PERMDISP, F=0.164404, p=0.660) and unweighted UniFrac (PERMANOVA, pseudo-F=4.31874, p=0.003, PERMDISP, F=8.00349, p=0.011) respectively.

P. agricolaris infected amoebae possess microbiomes that are significantly enriched in toxin degradation pathways

The consistent prevalence of *P. agricolaris* over the past two decades could be attributed to a variety of mutualistic effects, as seen in other systems. Therefore, we investigated other potential roles of *P. agricolaris* by comparing the natural microbiome functional potentials of infected and uninfected *D. discoideum*. We utilized PICRUSt2 to infer the metagenomic compositions of the *D. discoideum* microbiomes and used KEGG Orthology to assign and group functional pathways. We found that infected *D. discoideum* showed significant enrichments in pathways relating to Xenobiotic Biodegradation and Metabolism (Mann-Whitney/Kruskal-Wallis: $p = 0.0075$) (Figure 8). It must be noted, however, that these results are merely functional inferences, and further studies would be required to validate these findings.

Discussion

While previous studies of the *Paraburkholderia* – *D. discoideum* symbiosis have worked with samples isolated from nature, exploration of the farming phenomenon has been primarily laboratory based. Therefore, we aimed to investigate this interaction as it might occur in a more natural context. It has been previously established that *Paraburkholderia* is prevalent in natural *D. discoideum* populations (31). Here, we expound upon these findings by identifying *Paraburkholderia* prevalence in 2019 . Our survey of natural *D. discoideum* populations showed little to no incidences of *P. bonniea* and *P. hayleyella* infection, yet *P. agricolaris* remained prevalent. Furthermore, we found evidence that *P. agricolaris* is environmentally acquired in nature, which is supported by its high degree of genetic diversity. Overall, we did not find support for our hypothesis that the *Paraburkholderia* farming symbiont would increase the diversity of

D. discoideum microbiome. Rather, *P. agricolaris* tended to infectiously dominate the *D. discoideum* microbiome in nature. Artificial inoculation of each farming symbiont recapitulated this pattern, where *P. bonniea* and *P. hayleyella* also infectiously dominated the *D. discoideum* microbiome. While our results did not support an ecologically relevant enhancement of bacterial food carriage by *P. agricolaris*, our findings suggest a potential role in toxin degradation.

Differences in the soil environment and the laboratory environment may explain differences in the observation of farming. Laboratory tests of farming are generally in a controlled environment, starting with *D. discoideum* clones being passaged and cleared of natural transient bacteria and then provided with an ample supply of specific types of food bacteria (26). *Paraburkholderia* confers secondary food carriage under these conditions, mostly extracellular, and the extent of this carriage varies by *Paraburkholderia* genotype and food bacterial species (32,64). Although the mechanism is not clear, *Paraburkholderia* infection may inhibit the phagocytic pathway or otherwise disrupt the social amoeba immune system to allow food bacteria to persist. In the soil environment, many fruiting bodies contain a microbiome, both with and without *Paraburkholderia* (21,22). *Paraburkholderia* in our study did not increase the diversity of this microbiome indicative increasing secondary food carriage. Our natural population *D. discoideum* samples were passaged once, which likely decreased the microbiome diversity, especially when compared to our lab infection experiment where fruiting bodies were collected directly from the soil. In our laboratory infection experiment with its larger microbiome diversity, the overly abundant *Paraburkholderia* significantly

decreased the microbiome diversity. Perhaps *Paraburkholderia*-induced secondary food carriage, if it is happening in nature, is small in comparison to the natural microbiome. It is also possible that *P. agricolaris* competitively excludes other bacteria from infecting the amoeba fruiting body, just as the certain *Burkholderia* symbionts out compete non-native symbionts in the crypt of the bean bug *Riptortus pedestris* (65).

Though *Paraburkholderia* dominated the microbiome, it seldom accounted for 100% of microbiome composition. Thus, we cannot exclude the possibility that the remaining bacteria in *Paraburkholderia* infected microbiota are beneficial to the host. Although, if the presence of a bacterial microbiome in general is beneficial, uninfected amoebae would receive the same benefit, thus changing our understanding of farming. We did observe that *Paraburkholderia* infected amoebae had compositionally similar microbiomes, which could be consistent with farming if *Paraburkholderia* favored the carriage of particular bacteria more likely to be food. These compositional changes, however, were merely due to a decrease in phylogenetic diversity caused by *Paraburkholderia* infection. The only other significant effect of *Paraburkholderia* on the microbiome was its positive co-association with Rhizobiales. While the implications of this association are unclear, it merits further investigation, given observations of similar patterns in experimental studies (32) and the importance of Rhizobiales in the soil environment.

Although our findings did not support the idea that farming was the primary driver of *P. agricolaris* prevalence in nature, inference based metagenomic analyses suggest that *P. agricolaris* might play a role in toxin degradation for its amoeba host. This is consistent with previous demonstrations that *D. discoideum* spore production is significantly reduced when exposed to the toxin ethidium bromide but is not impacted under the same conditions when harboring *Paraburkholderia* (66). While this study suggested that farming associated bacteria over-compensated for fewer sentinel cells produced by *Paraburkholderia* infected hosts, our findings suggest that *P. agricolaris* itself could be filling this role. Our analyses only measured functional potential, not direct functional activity; however, such roles are assumed by *Burkholderia* symbionts of both insects and fungi. The bean bug *Riptortus pedestris* harbors *Burkholderia* symbionts that protect their host from the harmful effects of the insecticide fenitrothion, and other *Burkholderia* symbionts protect their fungal hosts from a variety of anti-fungal agents (29,67), yet others produce anti-fungal agents themselves (68).

There is conflicting evidence as to whether these *Paraburkholderia-D. discoideum* symbioses are co-evolved mutualisms. The prevalence of *P. agricolaris* in nature is consistent with beneficial effects that outweighs any cost of infection, as would be seen in a mutualism. Likewise, finding almost no incidence of *P. bonniea* or *P. hayleyella* infection could indicate their detriment or lack of relevance to host fitness. More time points would be needed to establish this trend ecologically. Indeed, some *P*. *hayleyella* have a demonstrably large fitness cost in laboratory assays, though *P. bonniea* has not been reported to significantly impact *D. discoideum* spore production (31,32). Many long

term co-evolved mutualisms lack horizontal transmission, and the symbiont genomes degrade as they adapt to living solely inside their host and vertical transmission (69). However, *P. agricolaris* possesses a larger genome than both *P. bonniea* and *P. hayleyella* (70). Due to its larger genome, *P. agricolaris* can utilize a wider diversity of carbon sources than its counterparts, which makes it more similar to other free-living *Paraburkholderia* and *Burkholderia* relatives. It also infects a fewer proportion of spores than the other *Paraburkholderia* symbiont species (32,64), suggesting that it is not strongly vertically transferred. *P. bonniea* and *P. hayleyella*, which infect a higher proportion of spores in laboratory experiments, may be more reliant on vertical transmission in nature as their smaller genomes suggest they may not be as capable of living freely in the soil.

Thus, it is possible that *P. agricolaris* is simply prevalent in *D. discoideum* populations because it is prevalent in, and subsequently acquired from, the soil environment. This is supported by the findings of Sallinger *et al.* (22), where *P. agricolaris* was detectable in relatively high frequency in bulk soil sample from which amoebae were cultured. Environmental acquisition could explain the high genetic diversity of *P. agricolaris* in natural populations, which differs from the pattern of a single beneficial haplotype sweeping through a population after initial infection, as often occurs in insect symbionts(71,72). Here, *P. agricolaris* would not be subjected to the population bottleneck and subsequent loss of genetic diversity elicited by vertical transmission. Additionally, a wider diversity of *P. agricolaris* haplotypes could be acquired from the soil upon initial infection. The intracellular and extracellular occupation of *P. agricolaris*

makes it both horizontally and vertically transmissible (32), a pattern has been observed in insects, where the oriental chinch bug obtains its *Burkholderia* symbiont from both parent and environment (73). It is also consistent with the observation of these same *Paraburkholderia* haplotypes infecting other, distantly related co-occurring species of social amoebae. This acquisition could be driven by either the host (74) or bacteria (75).The ease of environmental acquisition of *P. agricolaris* may ultimately lessen the pressure to develop and maintain vertical transmission and limit co-evolution. Such patterns have also been observed in symbioses between *Burkholderia* and broad-headed insects, where acquisition of *Burkholderia* from the environment dampens co-evolutionary dynamics (76,77).

Like *Paraburkholderia*, we found that Chlamydiae is highly prevalent in natural populations, though the reasons for this are less clear. Laboratory studies of Chlamydiae in *D. discoideum* found no fitness cost of infection (25), but potential benefits have yet to be explored. Like *Paraburkholderia*, Chlamydiae was overabundant in the microbiomes that it occupied, resulting in compositional uniformity. Co-association analyses within the microbiome showed positive associations between the order Chlamydiales and the orders Flavobacteriales and Peptostreptococcal, suggesting that this bacterium could be altering the social amoeba microbiome in meaningful ways. The high prevalence of Chlamydiae in natural populations suggests an ecological relevance for this symbiont that is worth further studying.

Overall, the variability regarding *P. agricolaris* localization, prevalence, transmission

patterns, genetic diversity, and costs of infection might indicate movement along the parasitism – mutualism spectrum. Such dynamics can be difficult to study in more complex systems. Thus, the *D. discoideum – Paraburkholderia* symbiosis, as well as the *D. discoideum* microbiome as a whole, offers a useful model for understanding symbiotic development. Future studies that aim to identify specific environmental parameters and soil conditions that either facilitate or hinder this association will be useful in providing further ecological contextualization. Furthermore, future studies that aim to investigate alternative relevance of *Paraburkholderia*, such as a role in toxin degradation, will further our understanding of the context dependent nature of symbiotic interactions in general.

Data Availability

All sequences used to confirm *Paraburkholderia* infection and identity in natural *D. discoideum* populations have been uploaded to GenBank under accession numbers MZ395652 - MZ395810. Sequences used for *P. agricolaris* haplotype analyses have been uploaded under accession numbers MZ399634 - MZ399696. Next-generation 16S rRNA sequences obtained for microbiome comparisons have been uploaded as BioProject PRJNA737516. Sequences used to confirm amoeba species identity were uploaded to GenBank under accession numbers MZ396209-MZ396243.

Acknowledgments

We would like to thank Julia Roberson for technical assistance with the laboratory inoculation experiment, Susanne DiSalvo for valuable discussions and comments on the manuscript, and Joan Strassmann for her feedback on this work. Funding was provided by the University of Central Arkansas (UCA) Advancement of Undergraduates in Research in the Science (AURS) program, the Arkansas Department of Higher Education Student Undergraduate Research Fellowship (SURF), and the UCA College of Natural Science and Mathematics. Bioinformatics were performed on the UCA Dell Blade Server System, "BRUCE," which was funded by NSF Award Number 1726009. This project has been supported by the Arkansas INBRE program, with a grant from the National Institute of General Medical Sciences (NIGMS), P20 GM103429 from the National Institute of Health.

References

1. Ewald PW. Transmission modes and evolution of the parasitism-mutualism continuum. Ann N Y Acad Sci. 1987;503:295–306.

2. Moran NA. Symbiosis. Curr Biol. 2006/10/24 ed. 2006 Oct 24;16(20):R866-71.

3. Anderson RM, May RM. Coevolution of hosts and parasites. Parasitology. 1982 Oct;85 (Pt 2):411–26.

4. Sun S, Noorian P, McDougald D. Dual Role of Mechanisms Involved in Resistance to Predation by Protozoa and Virulence to Humans. Front Microbiol. 2018;9:1017.

5. Dunn JD, Bosmani C, Barisch C, Raykov L, Lefrançois LH, Cardenal-Muñoz E, et al. Eat Prey, Live: Dictyostelium discoideum As a Model for Cell-Autonomous Defenses. Front Immunol [Internet]. 2018 [cited 2021 Jun 19];8. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2017.01906/full

6. Brüssow H. Bacteria between protists and phages: from antipredation strategies to the evolution of pathogenicity. Mol Microbiol. 2007 Aug;65(3):583–9.

7. Greub G, Raoult D. Microorganisms Resistant to Free-Living Amoebae. Clin Microbiol Rev. 2004;17(2):413–33.

608 8. Molmeret M, Horn M, Wagner M, Santic M, Abu Kwaik Y. Amoebae as training grounds for
609 intracellular bacterial pathogens. Appl Environ Microbiol. 2005 Jan;71(1):20–8.

610 9. Casadevall A. Evolution of intracellular pathogens. Annu Rev Microbiol. 2008/09/13 ed. 2008;62:19–
611 33.

612 10. Horn M, Wagner M. Bacterial endosymbionts of free-living amoebae. J Eukaryot Microbiol. 2004
613 Oct;51(5):509–14.

614 11. Schmitz-Esser S, Toenshoff ER, Haider S, Heinz E, Hoenninger VM, Wagner M, et al. Diversity of
615 bacterial endosymbionts of environmental acanthamoeba isolates. Appl Environ Microbiol. 2008
616 Sep;74(18):5822–31.

617 12. Delafont V, Samba-Louaka A, Bouchon D, Moulin L, Héchard Y. Shedding light on microbial dark
618 matter: a TM6 bacterium as natural endosymbiont of a free-living amoeba. Environ Microbiol Rep.
619 2015 Dec;7(6):970–8.

620 13. Horn M. Chlamydiae as symbionts in eukaryotes. Annu Rev Microbiol. 2008;62:113–31.

621 14. Okude M, Matsuo J, Nakamura S, Kawaguchi K, Hayashi Y, Sakai H, et al. Environmental chlamydiae
622 alter the growth speed and motility of host acanthamoebae. Microbes Environ. 2012;27(4):423–9.

623 15. Maita C, Matsushita M, Miyoshi M, Okubo T, Nakamura S, Matsuo J, et al. Amoebal endosymbiont
624 Neochlamydia protects host amoebae against Legionella pneumophila infection by preventing
625 Legionella entry. Microbes Infect. 2018 Apr;20(4):236–44.

626 16. Ishida K, Sekizuka T, Hayashida K, Matsuo J, Takeuchi F, Kuroda M, et al. Amoebal endosymbiont
627 Neochlamydia genome sequence illuminates the bacterial role in the defense of the host amoebae
628 against Legionella pneumophila. PloS One. 2014;9(4):e95166.

629 17. König L, Wentrup C, Schulz F, Wascher F, Escola S, Swanson MS, et al. Symbiont-Mediated Defense
630 against Legionella pneumophila in Amoebae. mBio. 2019 May 14;10(3):e00333-19.

631 18. Bozzaro S. The past, present and future of Dictyostelium as a model system. Int J Dev Biol.
632 2019;63(8-9–10):321–31.

633 19. Kessin RH. Dictyostelium: Evolution, Cell Biology, and the Development of Multicellularity.
634 Cambridge University Press; 2001. 318 p.

635 20. Chen G, Zhuchenko O, Kuspa A. Immune-like phagocyte activity in the social amoeba. Science.
636 2007/08/04 ed. 2007 Aug 3;317(5838):678–81.

637 21. Brock DA, Haselkorn TS, Garcia JR, Bashir U, Douglas TE, Galloway J, et al. Diversity of Free-Living
638 Environmental Bacteria and Their Interactions With a Bactivorous Amoeba. Front Cell Infect
639 Microbiol. 2018;8:411.

640 22. Sallinger E, Robeson MS, Haselkorn TS. Characterization of the bacterial microbiomes of social
641 amoebae and exploration of the roles of host and environment on microbiome composition.
642 Environ Microbiol. 2021 Jan;23(1):126–42.

643 23. Farinholt T, Dinh C, Kuspa A. Microbiome management in the social amoeba Dictyostelium
644 discoideum compared to humans. Int J Dev Biol. 2019;63(8-9–10):447–50.

645 24. Dinh C, Farinholt T, Hirose S, Zhuchenko O, Kuspa A. Lectins modulate the microbiota of social
646 amoebae. Science. 2018 Jul 27;361(6400):402–6.

647 25. Haselkorn TS, Jimenez D, Bashir U, Sallinger E, Queller DC, Strassmann JE, et al. Novel Chlamydiae
648 and Amoebophilus endosymbionts are prevalent in wild isolates of the model social amoeba
649 Dictyostelium discoideum. Environ Microbiol Rep [Internet]. [cited 2021 Jul 2];n/a(n/a). Available
650 from: https://sfamjournals.onlinelibrary.wiley.com/doi/abs/10.1111/1758-2229.12985

651 26. Brock DA, Douglas TE, Queller DC, Strassmann JE. Primitive agriculture in a social amoeba. Nature.
652 2011/01/21 ed. 2011 Jan 20;469(7330):393–6.

653 27. DiSalvo S, Haselkorn TS, Bashir U, Jimenez D, Brock DA, Queller DC, et al. Burkholderia bacteria
654 infectiously induce the proto-farming symbiosis of Dictyostelium amoebae and food bacteria. Proc
655 Natl Acad Sci U S A. 2015 Sep 8;112(36):E5029-5037.

656 28. Kaltenpoth M, Flórez LV. Versatile and Dynamic Symbioses Between Insects and Burkholderia
657 Bacteria. Annu Rev Entomol. 2020 Jan 7;65:145–70.

658 29. Nazir R, Tazetdinova DI, van Elsas JD. Burkholderia terrae BS001 migrates proficiently with diverse
659 fungal hosts through soil and provides protection from antifungal agents. Front Microbiol.
660 2014;5:598.

661 30. Esmaeel Q, Miotto L, Rondeau M, Leclère V, Clément C, Jacquard C, et al. Paraburkholderia
662 phytofirmans PsJN-Plants Interaction: From Perception to the Induced Mechanisms. Front
663 Microbiol. 2018;9:2093.

664 31. Haselkorn TS, DiSalvo S, Miller JW, Bashir U, Brock DA, Queller DC, et al. The specificity of
665 Burkholderia symbionts in the social amoeba farming symbiosis: Prevalence, species, genetic and
666 phenotypic diversity. Mol Ecol. 2019 Feb;28(4):847–62.

667 32. Khojandi N, Haselkorn TS, Eschbach MN, Naser RA, DiSalvo S. Intracellular Burkholderia Symbionts
668 induce extracellular secondary infections; driving diverse host outcomes that vary by genotype and
669 environment. ISME J. 2019 Aug;13(8):2068–81.

670 33. Oliver KM, Campos J, Moran NA, Hunter MS. Population dynamics of defensive symbionts in aphids.
671 Proc Biol Sci. 2008 Feb 7;275(1632):293–9.

672 34. Raper KB. The Dictyostelids. Princeton University Press; 2014. 467 p.

673 35. Baldauf SL, Romeralo M, Fiz-Palacios O, Heidari N. A Deep Hidden Diversity of Dictyostelia. Protist.
674 2018 Feb;169(1):64–78.

675 36. Salles JF, De Souza FA, van Elsas JD. Molecular method to assess the diversity of Burkholderia
676 species in environmental samples. Appl Environ Microbiol. 2002 Apr;68(4):1595–603.

677 37. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an
678 integrated and extendable desktop software platform for the organization and analysis of sequence
679 data. Bioinforma Oxf Engl. 2012 Jun 15;28(12):1647–9.

680 38. Spilker T, Baldwin A, Bumford A, Dowson CG, Mahenthiralingam E, LiPuma JJ. Expanded multilocus
681 sequence typing for burkholderia species. J Clin Microbiol. 2009 Aug;47(8):2607–10.

682 39. Ossewaarde JM, Meijer A. Molecular evidence for the existence of additional members of the order
683 Chlamydiales. Microbiol Read Engl. 1999 Feb;145 (Pt 2):411–7.

684 40. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for
685 Bigger Datasets. Mol Biol Evol. 2016 Jul;33(7):1870–4.

686 41. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al.
687 DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol Biol Evol. 2017 Dec
688 1;34(12):3299–302.

689 42. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: Detection and analysis of
690 recombination patterns in virus genomes. Virus Evol. 2015;1(1):vev003.

691 43. Redford AJ, Bowers RM, Knight R, Linhart Y, Fierer N. The ecology of the phyllosphere: geographic
692 and phylogenetic variability in the distribution of bacteria on tree leaves. Environ Microbiol. 2010
693 Nov;12(11):2885–93.

694 44. Sayers EW, Beck J, Bolton EE, Bourexis D, Brister JR, Canese K, et al. Database resources of the
695 National Center for Biotechnology Information. Nucleic Acids Res. 2021 Jan 8;49(D1):D10–7.

696 45. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution
697 sample inference from Illumina amplicon data. Nat Methods. 2016 Jul;13(7):581–3.

698 46. McDonald D, Clemente JC, Kuczynski J, Rideout JR, Stombaugh J, Wendel D, et al. The Biological
699 Observation Matrix (BIOM) format or: how I learned to stop worrying and love the ome-ome.
700 GigaScience. 2012 Jul 12;1(1):7.

701 47. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web
702 interface. Nucleic Acids Res. 2008 Jul 1;36(Web Server issue):W5-9.

703 48. Robeson MS, O'Rourke DR, Kaehler BD, Ziemski M, Dillon MR, Foster JT, et al. RESCRIPt:
704 Reproducible sequence taxonomy reference database management for the masses. bioRxiv. 2020
705 Oct 5;2020.10.05.326504.

706 49. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al. Optimizing taxonomic
707 classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin.
708 Microbiome. 2018 May 17;6(1):90.

709 50. Janssen S, McDonald D, Gonzalez A, Navas-Molina JA, Jiang L, Xu ZZ, et al. Phylogenetic Placement
710 of Exact Amplicon Sequences Improves Associations with Clinical Information. mSystems. 2018
711 Jun;3(3):e00021-18.

712 51. McKinney W. Data Structures for Statistical Computing in Python. In 2010.

713 52. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and
714 applications. BMC Bioinformatics. 2009 Dec 15;10:421.

715 53. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, et al. Normalization and microbial
716 differential abundance strategies depend upon data characteristics. Microbiome. 2017 Mar
717 3;5(1):27.

718 54. Pielou EC. The measurement of diversity in different types of biological collections. 1966;

719 55. Kruskal WH, Wallis WA. Use of Ranks in One-Criterion Variance Analysis. J Am Stat Assoc.
720 1952;47(260):583–621.

721 56. Martino C, Morton JT, Marotz CA, Thompson LR, Tripathi A, Knight R, et al. A Novel Sparse
722 Compositional Technique Reveals Microbial Perturbations. mSystems. 2019 Feb;4(1):e00016-19.

723 57. Lozupone C, Knight R. UniFrac: a new phylogenetic method for comparing microbial communities.
724 Appl Environ Microbiol. 2005 Dec;71(12):8228–35.

725 58. Lozupone CA, Hamady M, Kelley ST, Knight R. Quantitative and qualitative beta diversity measures
726 lead to different insights into factors that structure microbial communities. Appl Environ Microbiol.
727 2007 Mar;73(5):1576–85.

728 59. Faith DP, Minchin PR, Belbin L. Compositional Dissimilarity as a Robust Measure of Ecological
729 Distance. Vegetatio. 1987;69(1/3):57–68.

730 60. Chen J, Bittinger K, Charlson ES, Hoffmann C, Lewis J, Wu GD, et al. Associating microbiome
731 composition with environmental covariates using generalized UniFrac distances. Bioinforma Oxf
732 Engl. 2012 Aug 15;28(16):2106–13.

733 61. Chang Q, Luan Y, Sun F. Variance adjusted weighted UniFrac: a powerful beta diversity measure for
734 comparing communities based on phylogeny. BMC Bioinformatics. 2011 Apr 25;12:118.

735 62. Anderson MJ. A new method for non-parametric multivariate analysis of variance in ecology. In
736 2001.

737 63. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, et al. PICRUSt2 for prediction
738 of metagenome functions. Nat Biotechnol. 2020 Jun;38(6):685–8.

739 64. Miller JW, Bocke CR, Tresslar AR, Schniepp EM, DiSalvo S. Paraburkholderia Symbionts Display
740 Variable Infection Patterns That Are Not Predictive of Amoeba Host Outcomes. Genes. 2020 Jun
741 20;11(6):E674.

742 65. Itoh H, Jang S, Takeshita K, Ohbayashi T, Ohnishi N, Meng X-Y, et al. Host-symbiont specificity
743 determined by microbe-microbe competition in an insect gut. Proc Natl Acad Sci U S A. 2019 Nov
744 5;116(45):22673–82.

745 66. Brock DA, Callison WÉ, Strassmann JE, Queller DC. Sentinel cells, symbiotic bacteria and toxin
746 resistance in the social amoeba Dictyostelium discoideum. Proc Biol Sci. 2016 Apr
747 27;283(1829):20152727.

748 67. Kikuchi Y, Hayatsu M, Hosokawa T, Nagayama A, Tago K, Fukatsu T. Symbiont-mediated insecticide
749 resistance. Proc Natl Acad Sci U S A. 2012 May 29;109(22):8618–22.

750 68. Francoeur CB, May DS, Thairu MW, Hoang DQ, Panthofer O, Bugni TS, et al. Burkholderia from
751 Fungus Gardens of Fungus-Growing Ants Produces Antifungals That Inhibit the Specialized Parasite
752 Escovopsis. Appl Environ Microbiol. 2021 Jun 25;87(14):e0017821.

753 69. Moran NA. Accelerated evolution and Muller's rachet in endosymbiotic bacteria. Proc Natl Acad Sci
754 USA. 1996;93:2873–8.

755 70. Brock DA, Noh S, Hubert ANM, Haselkorn TS, DiSalvo S, Suess MK, et al. Endosymbiotic adaptations
756 in three new bacterial species associated with Dictyostelium discoideum: Paraburkholderia
757 agricolaris sp. nov., Paraburkholderia hayleyella sp. nov., and Paraburkholderia bonniea sp. nov.
758 PeerJ. 2020;8:e9151.

759 71. Himler AG, Adachi-Hagimori T, Bergen JE, Kozuch A, Kelly SE, Tabashnik BE, et al. Rapid spread of a
760 bacterial symbiont in an invasive whitefly is driven by fitness benefits and female bias. Science. 2011
761 Apr 8;332(6026):254–6.

762 72. Cockburn SN, Haselkorn TS, Hamilton PT, Landzberg E, Jaenike J, Perlman SJ. Dynamics of the
763 continent-wide spread of a Drosophila defensive symbiont. Ecol Lett. 2013 May;16(5):609–16.

764 73. Itoh H, Aita M, Nagayama A, Meng X-Y, Kamagata Y, Navarro R, et al. Evidence of environmental and
765 vertical transmission of Burkholderia symbionts in the oriental chinch bug, Cavelerius saccharivorus
766 (Heteroptera: Blissidae). Appl Environ Microbiol. 2014 Oct;80(19):5974–83.

767 74. Rashidi G, Ostrowski EA. Phagocyte chase behaviours: discrimination between Gram-negative and
768 Gram-positive bacteria by amoebae. Biol Lett. 2019 Jan 31;15(1):20180607.

769 75. Shu L, Zhang B, Queller DC, Strassmann JE. Burkholderia bacteria use chemotaxis to find social
770 amoeba Dictyostelium discoideum hosts. ISME J. 2018 Aug;12(8):1977–93.

771 76. Kikuchi Y, Hosokawa T, Fukatsu T. Insect-microbe mutualism without vertical transmission: a
772 stinkbug acquires a beneficial gut symbiont from the environment every generation. Appl Environ
773 Microbiol. 2007 Jul;73(13):4308–16.

774 77. Garcia JR, Laughton AM, Malik Z, Parker BJ, Trincot C, S L Chiang S, et al. Partner associations across
775 sympatric broad-headed bug species and their environmentally acquired bacterial symbionts. Mol
776 Ecol. 2014 Mar;23(6):1333–47.

777

Figure Legends:

Figure 1. *Paraburkholderia* and Chlamydiae are prevalent in natural *D. discoideum* populations. Pie charts indicate single and co-infection prevalence rates for each site sampled in the Mountain Lake, Virginia area.

Figure. 2. Phylogenetic evidence for environmental acquisition of *P. agricolaris.*
Maximum likelihood phylogenetic tree of *P. agricolaris* haplotypes constructed from the LepA housekeeping gene. Colored boxes represent every amoeba species found to contain the corresponding haplotype. Multiple amoebae species harboring the same haplotype was used as indication of environmental acquisition of *P. agricolaris.*

Figure. 3. *Paraburkholderia* infectiously dominates the *D. discoideum* microbiome. Bar plots show the members of the most abundant taxa within the microbiome of A) Naturally infected *D. discoideum* isolates and B) artificially infected *D. discoideum* isolates.

Fig. 4. *Paraburkholderia* does not diversity the microbiome. Alpha-diversity box plots showing the Pielou's evenness, observed ASVs, and Faith's phylogenetic diversity for A) isolates naturally infected with *Paraburkholderia* and B) isolates artificially infected with each *Paraburkholderia* symbiont. Boxes span the first and third quartiles and whiskers represent ± 1.5 IRQ. Letters shared in common denote no significant difference (Kruskal-Wallis α=0.05 after FDR correction.

Figure 5. High degrees of *P. agricolaris* abundance results in similar microbiome compositions among infected amoebae. Beta-diversity ordination plots comparing the microbiomes of *P. agricolaris* infected and uninfected *D. discoideum*, where each point represents an individual isolates microbiome. Kernel density estimate plots along the x and y axes indicate the distribution of points along the respective axes and highlight each groups concentration in ordination space. A) Weighted UniFrac PCoA plot showing the difference

between infected and uninfected microbiome structures. B) Unweighted UniFrac PCoA showing the difference in microbiome compositions. C) Aitchison PCA plot showing the contribution of key bacterial taxa to sample clustering.

Figure 6. Over-abundance of each *Paraburkholderia* species in the microbiome results in similar microbiome compositions among infected amoebae. Beta-diversity ordination plots comparing the microbiomes of *D. discoideum* artificially infected with each *Paraburkholderia* species to uninfected *D. discoideum*, where each point represents an individual isolates microbiome. Kernel density estimate plots along the x and y axes indicate the distribution of points along the respective axes and highlight each groups concentration in ordination space. A) Weighted UniFrac PCoA plot showing the difference between infected and uninfected microbiome structures. B) Unweighted UniFrac PCoA showing the difference in microbiome compositions. C) Aitchison PCA plot showing the contribution of key bacterial taxa to sample clustering.

Figure 7. *P. agricolaris* positively associates with members of the order Rhizobiales. Heatmap showing significant co-associations between bacterial orders within the natural *D. discoideum* microbiome. *P. agricolaris* is represented by the order Burkholderiales.

Figure 8. *P. agricolaris* infected microbiomes are significantly enriched in xenobiotic gene pathways due to high abundance of *P. agricolaris*. Box plot showing the abundance of gene pathways pertaining to xenobiotic degradation. Boxes span the first and third quartiles and whiskers represent ± 1.5 IRQ.

Figures

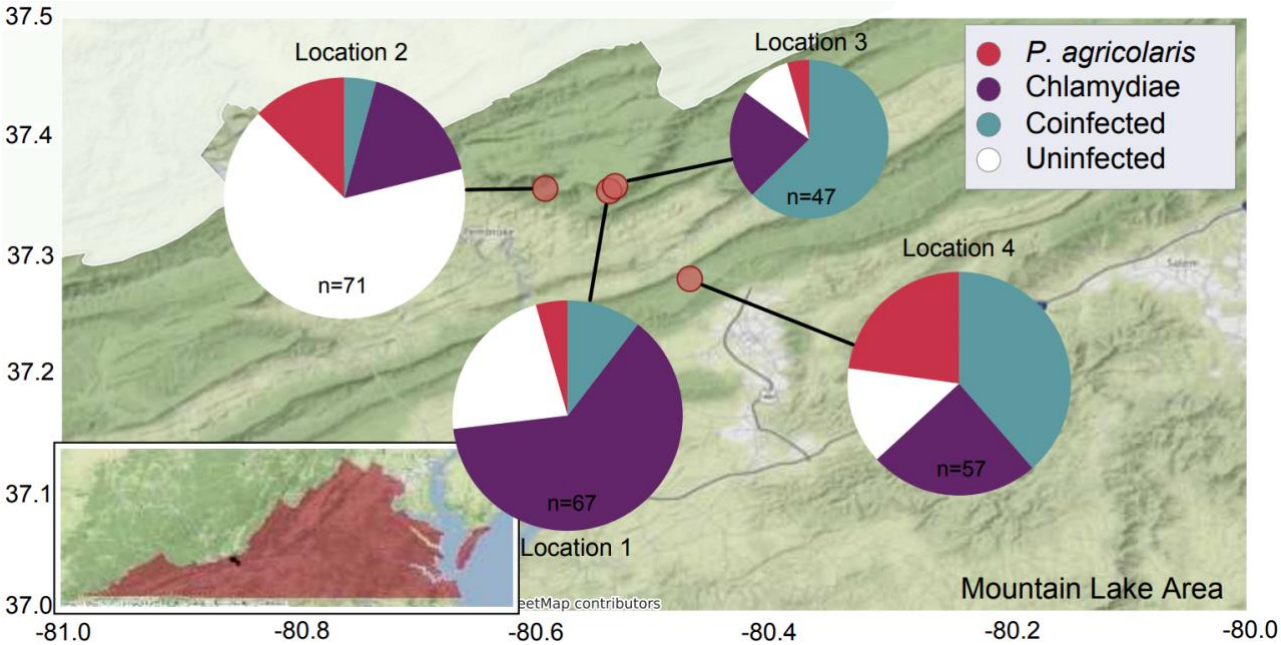

Fig. 1. *Paraburkholderia* and Chlamydiae prevalence and distribution in natural *D. discoideum* populations. Pie charts indicate single and co-infection prevalence rates for each site sampled in the Mountain Lake, Virginia area.

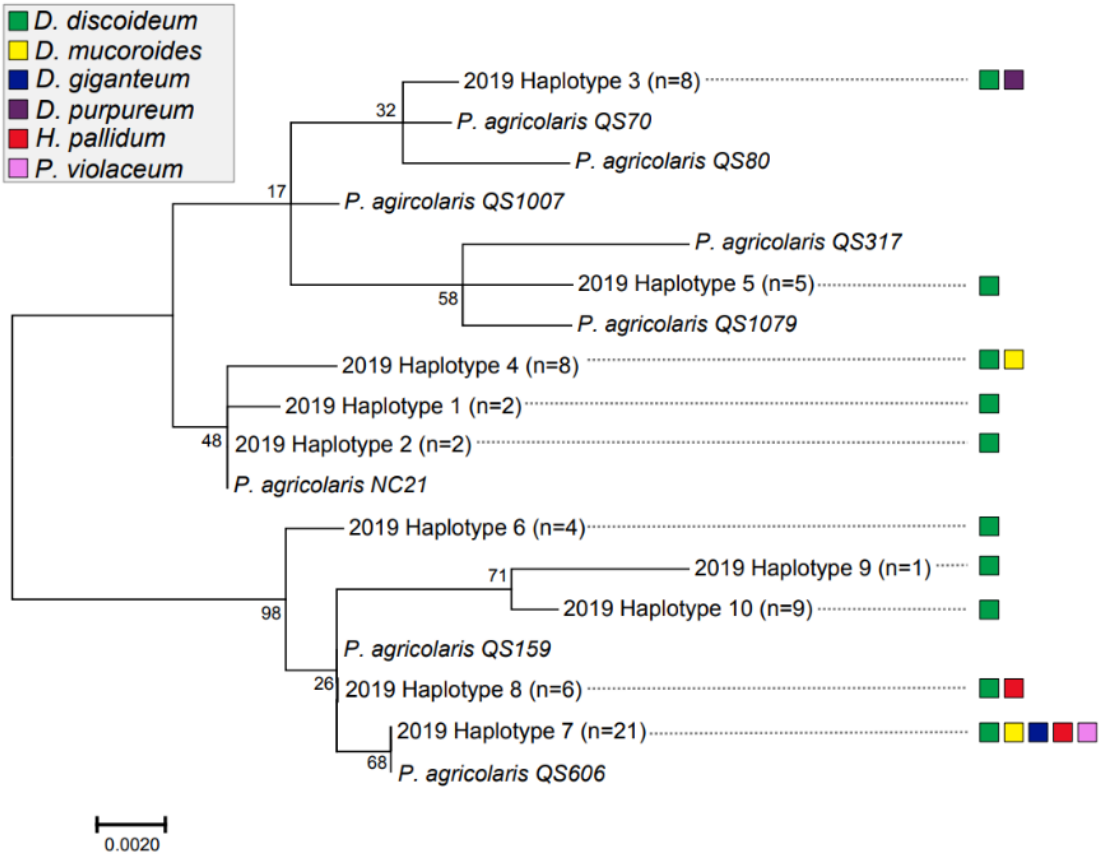

Fig. 2. A maximum likelihood phylogenetic tree of *P. agricolaris* haplotypes constructed from the LepA housekeeping gene. Colored boxes represent every amoeba species found to contain the corresponding haplotype. Multiple amoebae species harboring the same haplotype was used as indication of environmental acquisition of *P. agricolaris.*

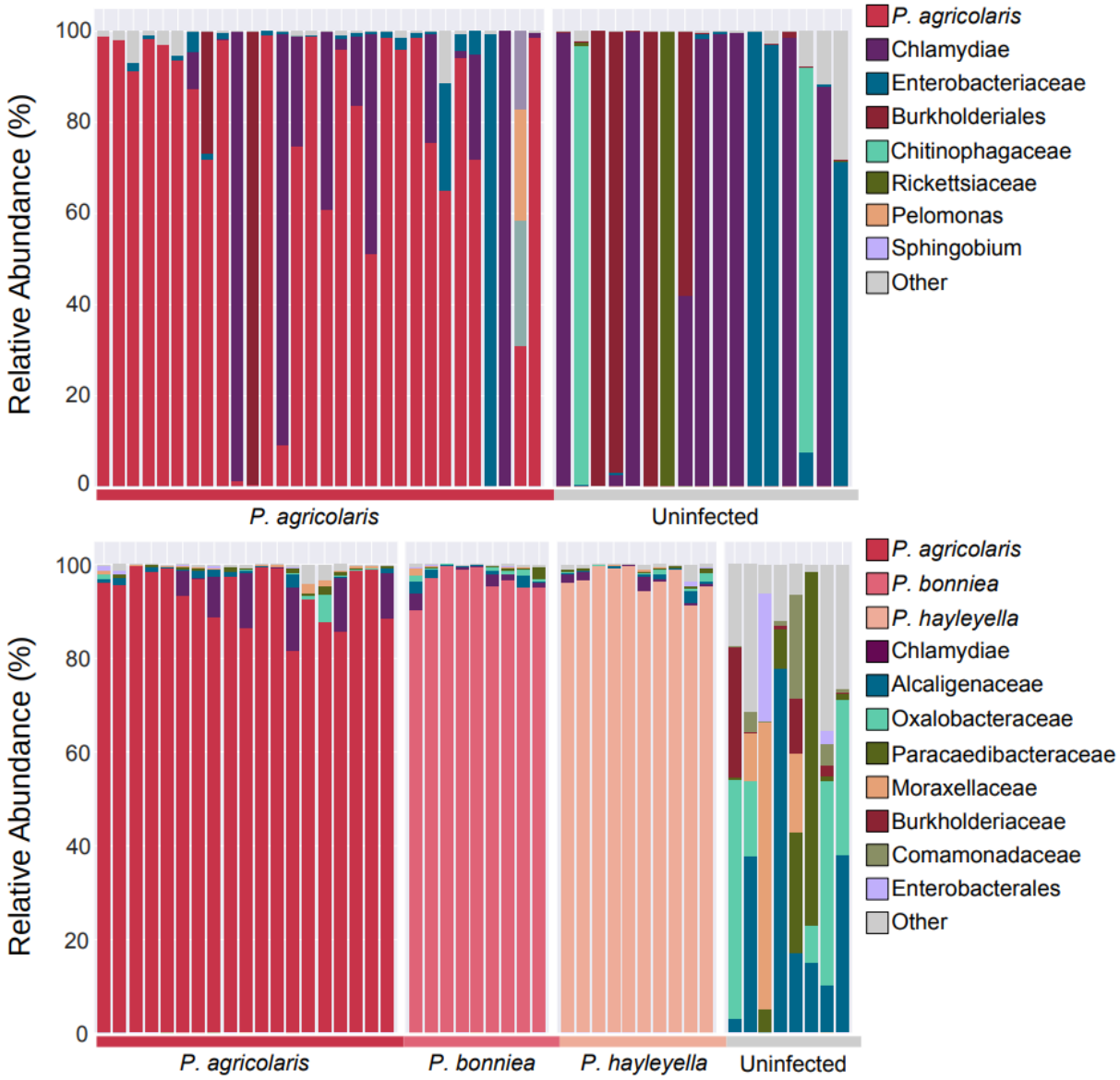

Fig. 3. *Paraburkholderia* infectiously dominates the *D. discoideum* microbiome. Bar plots show the members of the most abundant taxa within the microbiome of A) Naturally infected *D. discoideum* isolates and B) artificially infected *D. discoideum* isolates.

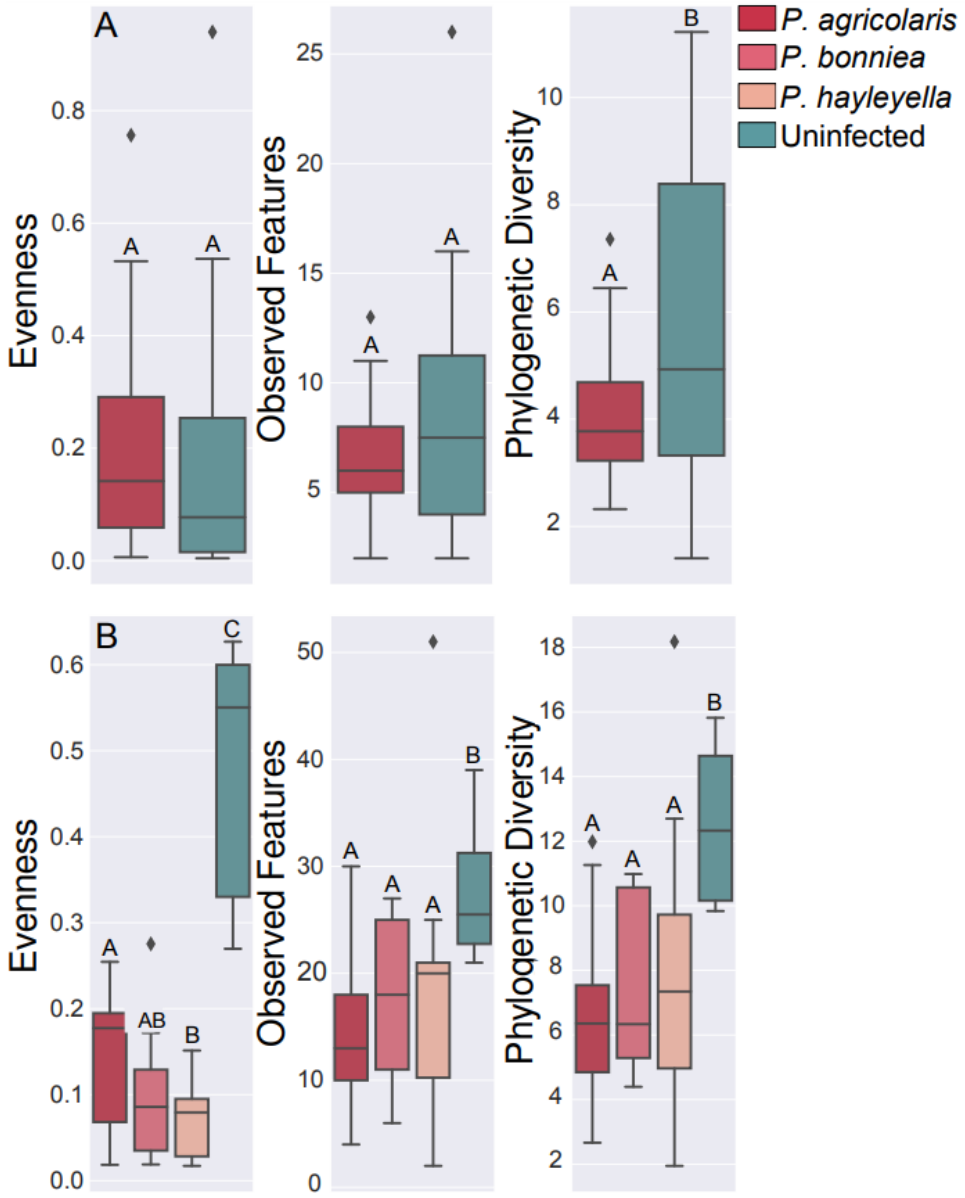

Fig. 4. Alpha-diversity box plots showing the Pielou's evenness, observed ASVs, and Faith's phylogenetic diversity for A) isolates naturally infected with *Paraburkholderia* and B) isolates artificially infected with each *Paraburkholderia* symbiont. Boxes span the first and third quartiles and whiskers represent ± 1.5 IRQ. Letters shared in common denote no significant difference (Kruskal-Wallis α=0.05 after FDR correction.

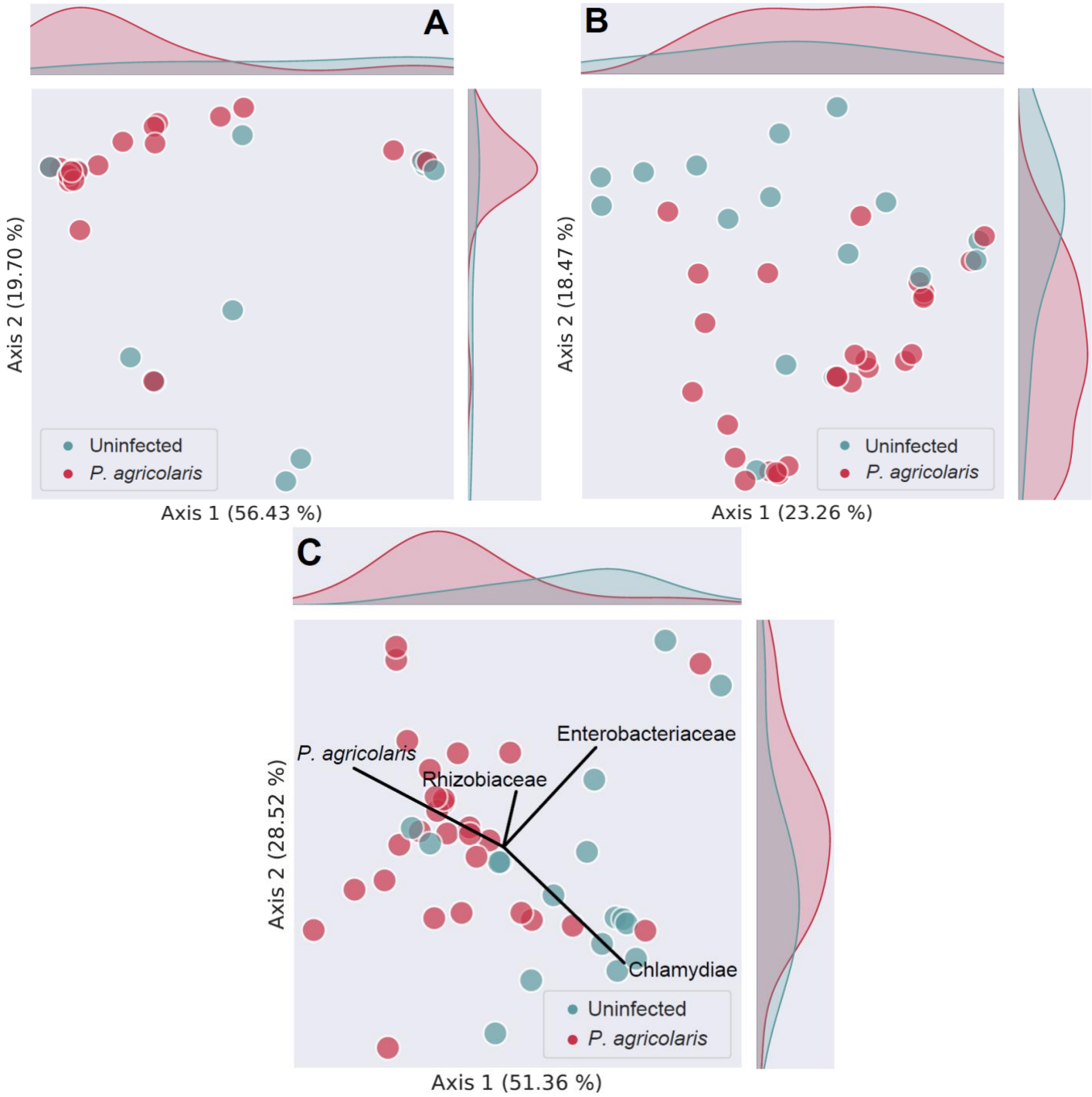

Fig. 5. Beta-diversity ordination plots comparing the microbiomes of *P. agricolaris* infected and uninfected *D. discoideum*, where each point represents an individual isolates microbiome. Kernel density estimate plots along the x and y axes indicate the distribution of points along the respective axes and highlight each groups concentration in ordination space. A) Weighted UniFrac PCoA plot showing the difference between infected and uninfected microbiome structures. B) Unweighted UniFrac PCoA showing the difference in microbiome compositions. C) Aitchison PCA plot showing the contribution of key bacterial taxa to sample clustering.

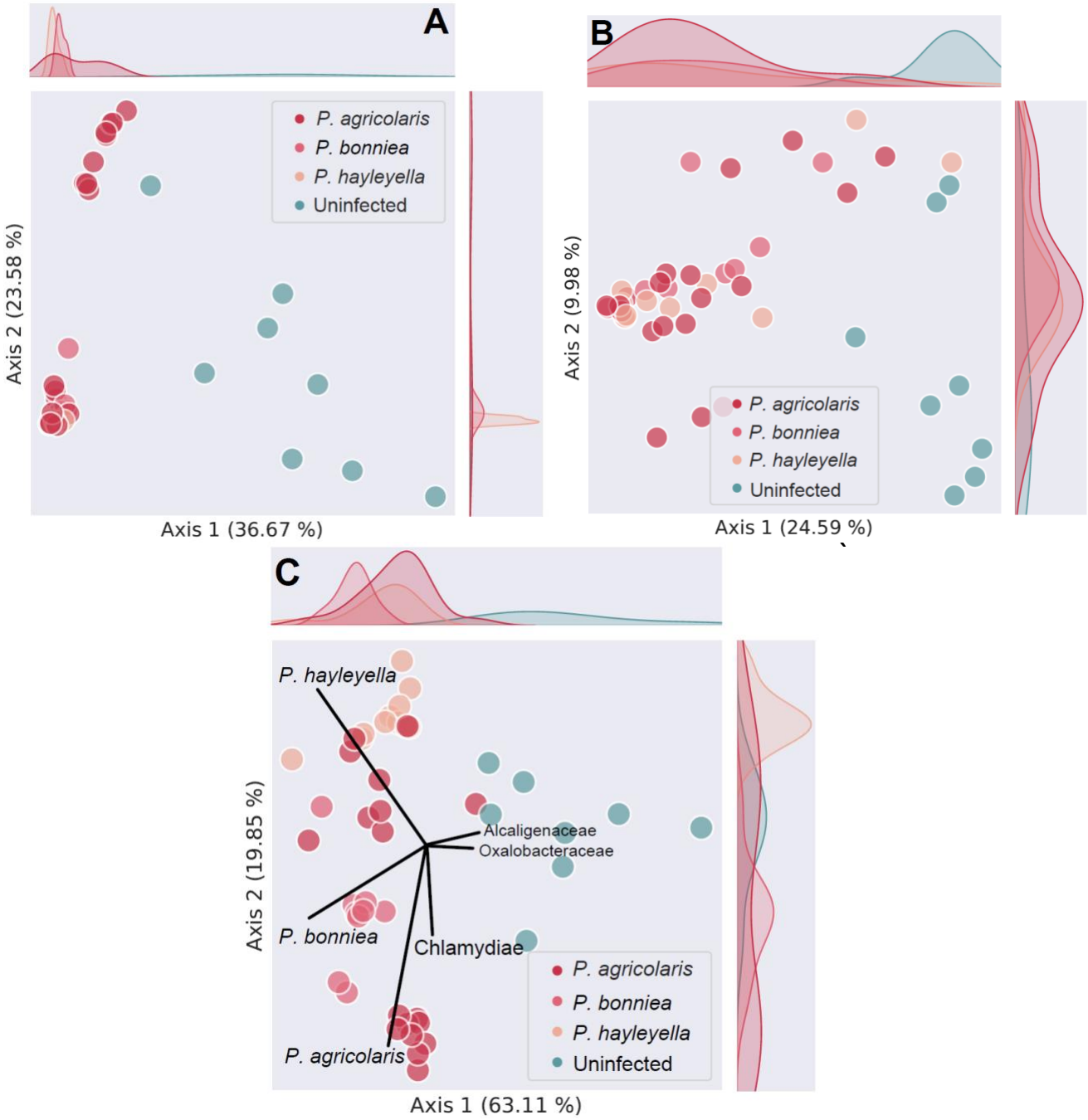

Fig. 6. Beta-diversity ordination plots comparing the microbiomes of *D. discoideum* artificially infected with each *Paraburkholderia* species to uninfected *D. discoideum*, where each point represents an individual isolates microbiome. Kernel density estimate plots along the x and y axes indicate the distribution of points along the respective axes and highlight each groups concentration in ordination space. A) Weighted UniFrac PCoA plot showing the difference between infected and uninfected microbiome structures. B) Unweighted UniFrac PCoA showing the difference in microbiome compositions. C) Aitchison PCA plot showing the contribution of key bacterial taxa to sample clustering.

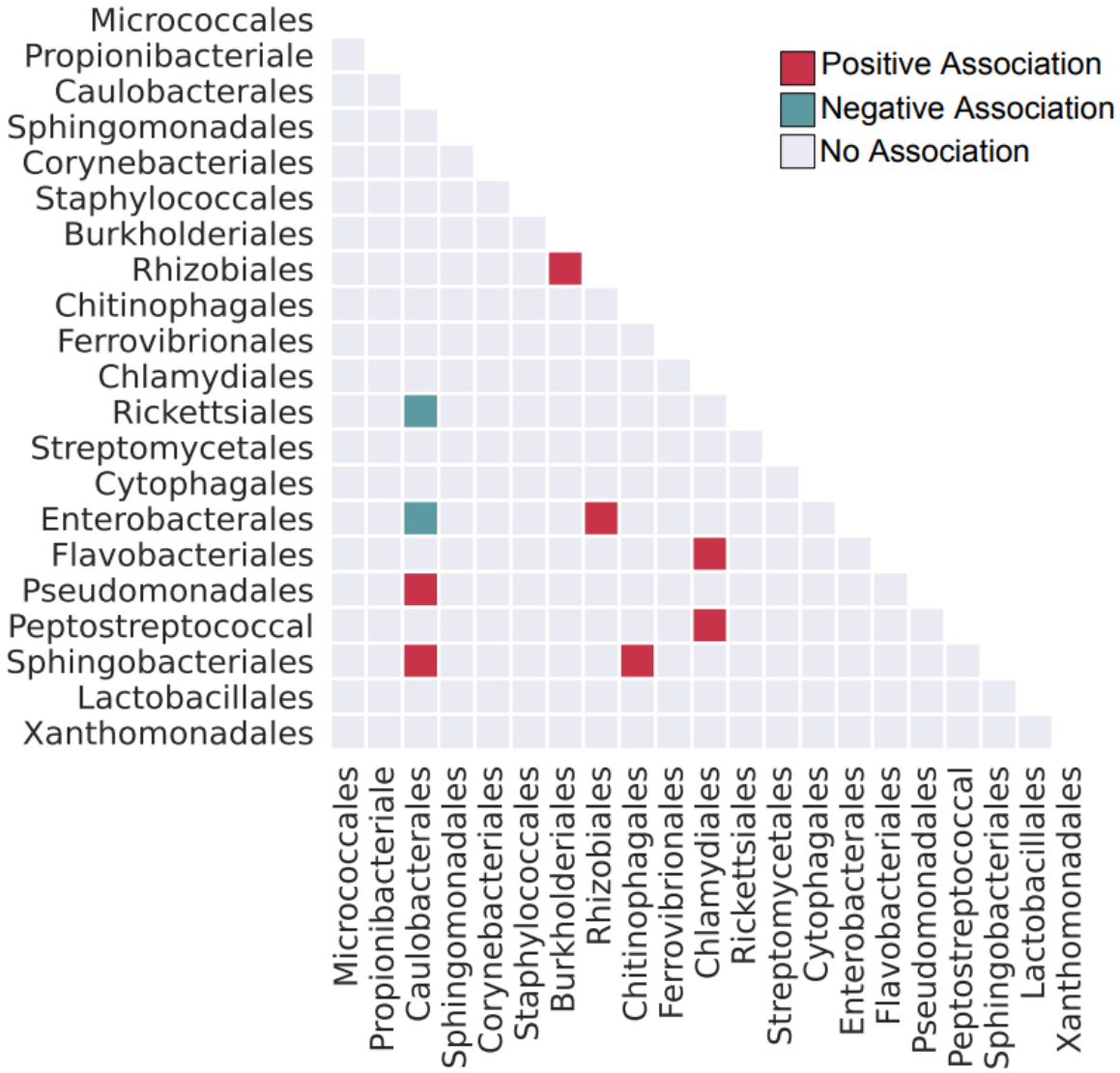

Fig. 7. Heatmap showing significant co-associations between bacterial orders within the natural *D. discoideum* microbiome. *P. agricolaris* is represented by the order Burkholderiales, and positively associates with members of the order Rhizobiales.

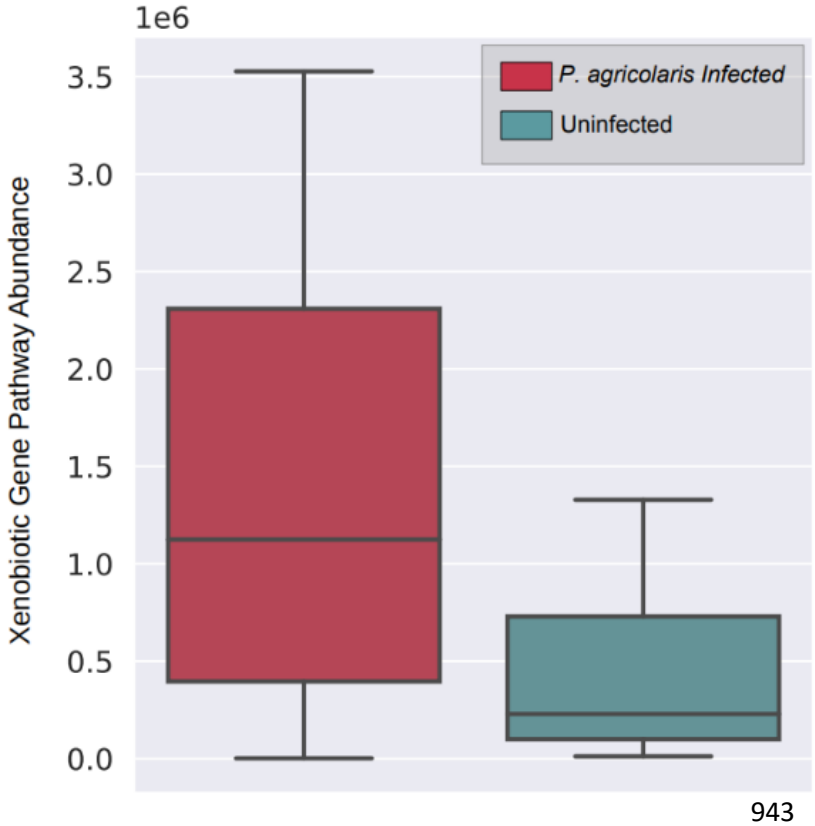

Fig. 8. Box plot showing the abundance of gene pathways pertaining to xenobiotic degradation. Boxes span the first and third quartiles and whiskers represent ± 1.5 IRQ. *P. agricolaris* infected microbiomes are significantly enriched in xenobiotic gene pathways due to high abundance of *P. agricolaris*.

Supplemental Tables

Supplemental Table 1. Soil sample metadata

Soil Sample	Location Name	GPS Coordinates	Description	Soil Temperature °C
Vial Sample 1	Blueberry Ridge Trail	N 37.357 W 80.359	Under log w/ foliage	16.5
Vial Sample 2	Indian Trail	N 37.357 W 80.534	Under foliage	14.6
Vial Sample 3	Indian Trail	N 37.358 W 80.533	Under and within rotting log	15
Vial Sample 4	Indian Trail	N 37.358 W 80.533	Under and within rotting log	14.5
Vial Sample 5	Indian Trail	N 37.357 W 80.534	Inside bowl of rotting stump	17
Vial Sample 6	Bald Knob Trail	N 37.354 W 80.537	Under log	14.5
Vial Sample 7	Bald Knob Trail	N 37.354 W 80.538	Inside of log	16
Bag Sample 1	Bald Knob Trail	N 37.353 W 80.537	Under log	17
Bag Sample 2	Bald Knob Trail	N 37.353 W 80.538	Under ground foliage	15.1
Bag Sample 3	Rocky Sink Road	N 37. 339 W 80.505	Under ground foliage	17
Bag Sample 4	Ricky Sink Road	N 37.340 W 80. 506	Under ground foliage	20
Bag Sample 5	Cougar Lane	N 37.356 W 80.464	Under ground foliage	17.2
Vial Sample 8	Cougar Lane	N 37.356 W 80.464	Under ground foliage	17.2
Vial Sample 9	Clover Hollow Road	N 37.315 W 80.497	Inside of rotting stump	N/A
Bag Sample 6	Cascade Falls	N 37.353 W 80.598	Under log	16.2
Bag Sample 7	Cascade Falls	N 37.354 W 80.598	Under ground foliage near log	17
Vial Sample 10	Cascade Falls	N 37.355 W 80.595	Under log, no foliage	18
Vial Sample 11	Cascade Falls	N 37.355 W 80.595	In log, no foliage	19.4
Vial Sample 12	Cascade Falls	N 37.356 W 80.594	Under log and foliage, near water	18.5
Vial Sample 13	Cascade Falls	N 37.356 W 80.592	Within divot in living tree	18.5
Bag Sample 8	Cascade Falls	N 37.356 W 80.592	Under ground foliage near log	16.1
Vial Sample 14	Cascade Falls	N 37.358 W 80.589	Under log and foliage	14.5
Vial Sample 15	Cascade Falls	N 37.359 W 80.588	Under log and foliage	14.5
Bag Sample 9	Cascade Falls	N 37.359 W 80.588	Under log and foliage	14.5
Vial Sample 16	Cascade Falls	N 37.361 W 80.586	Beside rotting stump	17.5
Bag Sample 10	Cascade Falls	N 37.361 W 80.586	Next to rotting log, under foliage, ontop of rocks	16.6
Vial Sample 17	Cascade Falls	N 37.361 W 80.586	Under rotting log and foliage	19
Bag Sample 11	Pandapas	N 37.281 W 80.470	Under foliage with pine mixed in	17.8
Vial Sample 18	Pandapas	N 37.281 W 80.470	Under foliage with pine mixed in	17.8
Vial Sample 19	Pandapas	N 37.281 W 80.470	Under rotting tree near water	20.3
Vial Sample 20	Pandapas	N 37.280 W 80.470	Under foliage, many roots present	16.5
Vial Sample 21	Pandapas	N 37.280 W 80.467	Between multiple decomposing logs, under foliage	18.5
Bag Sample 12	Pandapas	N 37.282 W 80.465	Between multiple decomposing logs, under foliage	18
Vial Sample 22	Pandapas	N 37.280 W 80.467	Between multiple decomposing logs, under foliage	18.5
Bag Sample 13	Pandapas	N 37.280 W 80.472	Beside rotting log, under foliage, many roots present	16.5
Vial Sample 23	Pandapas	N 37.280 W 80.472	Inside of rotting stump	17.5

Supplemental Table 2. Microbiological media preparation and KK2 extraction buffer

Hay agar ingredients (per liter)	Hay agar preparation protocol
15 g hay	1. Boil 15 g of hay in 1.5 L of DI H2O
1.5 L deionized H2O	2. Filter out hay from solution
1.5 g KH2PO4	3. Add each ingredient to hay infused solution
0.62 g Na2HPO4	4. Autoclave hay infused solution
15 g agar	5. Pipette 30 mL of solution into 100 mm petri dish

SM/5 agar ingredients (per liter)	SM/5 agar preparation protocol
2 g glucose	1. Mix each ingredient into 1 L of deionized water
2 g BactoPeptone	2. Autoclave solution
2 g yeast extract	3. Pipette 30 mL of solution into 100 mm petri dish
0.2 g MgCl2	
1.9 g KHPO4	
1 g KH2PO4	
15 g agar	

KK2 Spore Buffer ingredient (per liter)
2.25 g KH2PO4
0.67 g K2HPO4
1.0 L deionized water

963

Supplemental Table 3. Primer sets

Primer	Sequence	Reference
D307F	5'-GTTTGGCCTACCATGGTTGTAA-3'	Baldauf *et. al.,* 2018
D862R	5'-GGGSGTTCATATTGGGGCG-3'	Baldauf *et. al.,* 2018
Burk3F	5'-CTGCGAAAGCCGGAT-3'	Salles *et. al.,* 2002
Burk3R	5'-TGCCATACTCTAGCYYGC-3'	Salles *et. al.,* 2002
LepAF	5'-CTSATCATCGAYTCSTGGTTCG-3'	Spilker *et. al.,* 2009
LepAR	5'-CGRTATTCCTTGAACTCGTARTCC-3'	Spilker *et. al.,* 2009
CHL16SFOR2	5'-CGTGGATGAGGCATGCAAGTCGA-3'	Ossewaarde and Meijer, 1999
CHL16SREV2	5'-CAATCTCTCAATCCGCCTAGACGTCTTAG-3'	Ossewaarde and Meijer, 1999
AmoF	5'-AAAACCCTTGCGAGGGTTCTTGAAG-3'	Haselkorn *et. al.,* 2020
AmoR	5'-GCCTTCGCTGTTGGTATTCCTTATGG-3'	Haselkorn *et. al.,* 2020
799F	5'-AACMGGATTAGATACCCKG-3'	Chelius & Triplett, 2001
1115R	5'-AGGGTTGCGCTCGTTG-3'	Reysenbach & Pace, 1995

970

971

972

973

974

975

Supplemental Table 4. Prevalence, co-associations, and genetic community comparisons of symbiont screenings

Location	Location name	*P. agricolaris* prevalence (%)
1	Bald Knob Trail	14.93
2	Cascade Falls	16.9
3	Indian Trail	59.57
4	Pandapas	61.4
5	Total	35.12

Bald Knob Trail

Fisher Exact Test for Coassociation	Observed Coinfections	Expected Coinfections
P. agricolaris - Chlamydiae	7	7.313
P. agricolaris - Amoebophilus	5	3.134
Chlamydiae - Amoebophilus	16	15.36

Cascade Falls

Fisher Exact Test for Coassociation	Observed Coinfections	Expected Coinfections
P. agricolaris - Chlamydiae	3	2.54
P. agricolaris - Amoebophilus	2	1.35
Chlamydiae - Amoebophilus	3	2.11

Indian Trail

Fisher Exact Test for Coassociation	Observed Coinfections	Expected Coinfections
P. agricolaris - Chlamydiae	26	23.23
P. agricolaris - Amoebophilus	5	4.17
Chlamydiae - Amoebophilus	6	5.809

Pandapas

Fisher Exact Test for Coassociation	Observed Coinfections	Expected Coinfections
P. agricolaris - Chlamydiae	22	22.11
P. agricolaris - Amoebophilus	5	3.684
Chlamydiae - Amoebophilus	5	3.789

P. agricolaris populations Fst values	Column1	Column2
Population 1	**Population 2**	**Fst**
Bald Knob Trail	Cascade Falls	0.68667
Bald Knob Trail	Indian Trail	0.88679
Bald Knob Trail	Pandapas	0.87908
Cascade Falls	Indian Trail	0.13982
Cascade Falls	Pandapas	0.6532
Indian Trail	Pandapas	0.77522

Supplemental Table 5. Soil samples and *P. agricolaris* haplotypes isolated from them

Location	Soil Sample Code	*P. agricolaris* haplotypes present
Bald Knob Trail	1B	10
Cascade Falls	2B1	5,6,7
Indian Trail	3A1	5,6,7
Indian Trail	3A2	7
Indian Trail	3B	7,8
Pandapas	4A1	1,2,3,4,5
Pandapas	4A2	1,3,4
Pandapas	4A3	4