



## RESEARCH

e-Xtra\*

# The Foliar Microbiome Suggests that Fungal and Bacterial Agents May be Involved in the Beech Leaf Disease Pathosystem

Carrie J. Ewing,<sup>1,†</sup> Jason Slot,<sup>1</sup> María-Soledad Benítez,<sup>1</sup> Cristina Rosa,<sup>2</sup> Antonino Malacrinò,<sup>3</sup> Alison Bennett,<sup>3</sup> and Enrico Bonello<sup>1</sup>

<sup>1</sup> Department of Plant Pathology, The Ohio State University, Columbus, OH 43210

<sup>2</sup> Department of Plant Pathology and Environmental Microbiology, The Pennsylvania State University, State College, PA 16801

<sup>3</sup> Department of Evolution, Ecology, and Organismal Biology, The Ohio State University, Columbus, OH 43210

Accepted for publication 30 January 2021.

## ABSTRACT

Beech leaf disease (BLD) is a recently discovered disease that is causing severe damage to American beech (*Fagus grandifolia*) in northeastern North America. The recently described nematode *Litylenchus crenatae* subsp. *mccannii* was detected in BLD-affected foliage and may be associated with the disease. However, speculation on the direct role of the nematode in infection still remains. In this study, we profiled the microbial communities associated with asymptomatic, symptomatic, and naïve (control) American beech foliage by using a high-throughput sequence-based metabarcoding analysis of fungi, bacteria, phytoplasmas, and nematodes. We then used both a differential abundance analysis and indicator species analysis as well as several diversity metrics to try and discover microbes associated only with symptomatic foliage. To do so, we amplified the organism-specific phylogenetic informative regions of the 16S, 18S, and internal transcribed

spacer (ITS)1 regions using Illumina MiSeq. Our results detected the amplicon sequence variant (ASV) associated with the nematode *L. crenatae* subsp. *mccannii* but in all symptom types. However, four ASVs associated with the bacterial genera *Wolbachia*, *Erwinia*, *Paenibacillus*, and *Pseudomonas* and one ASV associated with the fungal genus *Paraphaeosphaeria* were detected only in symptomatic samples. In addition, we identified significant differences based on symptom type in both the  $\alpha$ - and  $\beta$ -diversity indices for the bacterial and fungal communities. These results suggest that *L. crenatae* subsp. *mccannii* may not be fully responsible for BLD but, rather, that other microbes may be contributing to the syndrome, including the putative nematode endosymbiont *Wolbachia* sp.

**Keywords:** bacteriology, endophytes, microbiome, mycology, nematology, plant pathology

<sup>†</sup>Corresponding author: C. J. Ewing; ewing.244@buckeyemail.osu.edu

**Author contributions:** P.B., J.S., and C.J.E. conceived and designed experiments. C.J.E. performed experiments and analyzed data. J.S., M.B., C.R., A.M., and A.B. contributed reagents, materials, and analytical tools. C.J.E. wrote initial manuscript. C.J.E., P.B., J.S., M.B., C.R., and A.M. revised the manuscript. All authors approved the final manuscript.

**Funding:** Funding for this project was provided by United States Department of Agriculture National Institute of Food and Agriculture Exploratory Research Program (award number 20176703026610) and by state and federal funds appropriated to the Ohio Agricultural Research and Development Center, Ohio State University, College of Food, Agricultural, and Environmental Sciences.

\*The e-Xtra logo stands for “electronic extra” and indicates that supplementary materials are published online.

The author(s) declare no conflict of interest.

In the spring of 2012, American beech (*Fagus grandifolia* Ehrh.) trees located in the northeastern portion of Ohio in Lake County Metroparks began showing symptoms of an unidentified disease, now known as beech leaf disease (BLD) (Ewing et al. 2019). The symptoms of the disease manifest as either an interveinal darkening of the leaf (hereafter referred to as banding) or deformation, darkening, and thickening of the leaf (hereafter referred to as crinkling) (Ewing et al. 2019). These symptoms are produced independently at bud break and do not change throughout the growing season (Daniel Volk and colleagues at Cleveland Metroparks, unpublished and personal communication). Eventually, buds cease to produce new leaves in the succeeding growing season, leading to tree mortality. Since its discovery in 2012, the disease has spread to American beech in Pennsylvania, Connecticut (Marra and LaMondia 2020), New York, and the province of Ontario, Canada (Supplementary Fig. S1). Recently, the disease has also been reported in Massachusetts and Rhode Island (Nick Brazee, University of Massachusetts, personal communication). The disease has also affected other beech species in nursery and arboretum settings, including

European (*F. sylvatica* L.) and oriental (*F. orientalis* Lipsky) beech (Ewing et al. 2019).

The symptoms of BLD do not resemble any other currently known diseases of American beech trees, which led us to believe that this is a novel disease caused by an unknown pest or pathogen. Carta et al. (2020) speculated that the interveinal darkening of the leaves was similar to damage caused by eriophyid mites, and the angular leaf spots within the darkened areas were similar to damage from foliar nematodes. They determined that a new subspecies of nematode, *Litylenchus crenatae* subsp. *mccannii* (LCM), was present in the leaves, and nematodes of this species collected from infested leaves produced symptoms of BLD when inoculated into the buds of healthy American beech trees (Carta et al. 2020). However, because the nematodes they used for inoculations were collected from symptomatic trees, the question remains as to whether the nematode is the sole causal agent of the disease or if other pathogens, associated with the nematode, may be necessary for full disease development.

In a recent study by Burke et al. (2020), terminal restriction fragment length polymorphism (T-RFLP) was used to examine the microbial communities of American beech foliage and better understand the role of the nematode in BLD infections. Their study determined that 36% of asymptomatic American beech leaves contained LCM, suggesting that LCM may not be the direct cause of BLD (Burke et al. 2020). By cloning and sequencing gene marker fragments, the study also showed that bacterial communities on leaves were different in the presence of BLD symptoms and that some bacterial genera, including *Wolbachia* and *Mucilaginibacter*, may be involved in BLD progression (Burke et al. 2020). This study paints a broader picture of the role that microorganisms could be having in BLD infections but does not confirm that those microorganisms are the causal agent of the disease.

We hypothesized that a more in-depth, high-throughput sequencing-based metabarcoding analysis of the microbial communities present in or on symptomatic American beech leaves compared with asymptomatic leaves would point to the true causal agent of BLD and either confirm or conflict with the LCM hypothesis. Such an approach would also contribute to a richer understanding of the beech foliar phytobiome than can be achieved with T-RFLP analysis.

## MATERIALS AND METHODS

**Foliage sampling.** In 2017, with assistance from members of Cleveland Metroparks and the Pennsylvania Bureau of Forestry, we established eight permanent sampling plots (Supplementary Fig. S1). Four plots were located in the general area of infestation in northeast Ohio, and two of those plots contained 10 symptomatic trees each and two contained 10 asymptomatic trees each. Similarly, three plots were created in the general area of infestation in northwest Pennsylvania, and two of those plots contained 10 symptomatic trees each and one contained 20 asymptomatic trees. One plot was established in central Ohio, well outside of the area of infestation, and contained 20 trees with no BLD symptoms, which we defined as the naïve symptom type. In 2018, leaves were sampled from all trees at two different time points (i.e., at full leaf expansion in May and then later in August). Samples were immediately flash frozen and stored in liquid nitrogen until arrival at the lab, where they were stored at  $-80^{\circ}\text{C}$ . Four to six leaves were collected from each tree but, for the symptomatic trees, four to six banded leaves and four to six crinkled leaves were collected as separate samples. Asymptomatic leaves were not collected from symptomatic trees because we assumed that all leaves on an infected tree

may harbor the causal agent. All samples were separated based on time point (May and August), symptom type (banding, crinkling, asymptomatic, and naïve), and location (northwest Pennsylvania, northeast Ohio, and central Ohio), creating 14 different sample types (Supplementary Table S1). However, the asymptomatic site in Pennsylvania could not be sampled in May; therefore, only 13 sample combinations were used.

**Amplicon library construction and sequencing.** The foliage samples were ground in liquid nitrogen, and total DNA was extracted from 50 mg of the ground tissue using Qiagen's DNeasy PowerPlant Pro Kit (Qiagen, Germantown, MD, U.S.A.) following the manufacturer's instructions for difficult samples. DNA was extracted from each individual sample (each comprising four to six leaves), then pooled by combining the DNA from five trees of each of the 13 sample types to establish four biological replicates per sample type (e.g., DNA from five randomly selected trees from the two Ohio asymptomatic sites sampled in May would make up one biological replicate). This resulted in 52 biological replicates (Supplementary Table S2). We chose to pool DNA instead of biomass to ensure that the DNA from the five trees was equally represented across the replicates. The quality of the extracted DNA was tested using a NanoDrop UV 1000 Visible spectrophotometer (Thermo Fisher Scientific, Waltham, MA, U.S.A.).

To characterize epiphytic and endophytic eukaryotic (specifically nematode), phytoplasm, bacterial, and fungal communities present in the leaves, each sample was used as a template for PCR to amplify the phylogenetically informative regions of 18S (nematodes), 16S (bacteria and phytoplasm), and the internal transcribed spacer (ITS)1 region (fungi). All primer sequences are available in the supplementary material (Supplementary Table S3). All group-specific PCR products were submitted for library construction and sequencing at the Ohio Agricultural Research and Development Center Molecular and Cellular Imaging Center, where they were indexed, pooled, and cleaned prior to sequencing.

To survey nematode communities, a seminested PCR approach was conducted using primers NemF, NF1, and 18Sr2b modified to contain Illumina overhang adapters (Sapkota and Nicolaisen 2015). The first PCR assays were carried out in 25- $\mu\text{l}$  volumes using Thermo Scientific's *Taq* DNA Recombinant (Thermo Fisher Scientific) and contained 1 $\times$  PCR buffer, 1.5 mM  $\text{MgCl}_2$ , 0.2 mM dNTPs, 1  $\mu\text{M}$  each NemF and 18Sr2b primers, 1 U of *Taq* Polymerase, and 1 ng of template DNA (Sapkota and Nicolaisen 2015). Following the protocol as detailed by Sapkota and Nicolaisen (2015), the reactions were amplified using a PTC-0200 DNA Engine thermocycler (Bio-Rad, Hercules, CA, U.S.A.) at the following conditions: initial denaturation at  $94^{\circ}\text{C}$  for 5 min; followed by 20 cycles of denaturation at  $94^{\circ}\text{C}$  for 30 s, annealing at  $53^{\circ}\text{C}$  for 30 s, and extension at  $72^{\circ}\text{C}$  for 1 min; with a final extension performed at  $72^{\circ}\text{C}$  for 10 min. The PCR product was diluted 1:10 and used as the template in the second PCR along with the NF1 and 18S2b primers. The remaining reaction reagents and conditions were the same as the first PCR, except the annealing temperature was raised to  $58^{\circ}\text{C}$  for the second PCR. The final expected band size was 420 bp.

To detect phytoplasm, a nested PCR using phytoplasm-specific primers was used to amplify the 16S ribosomal RNA (rRNA) region. The first PCR used primers P1 and TINT (Smart et al. 1996) and the second PCR used primers R16F2n and R16R2 (Gundersen and Lee 1996). Based on a modified protocol by Gundersen and Lee (1996), the first PCR assays were carried out in 25- $\mu\text{l}$  volumes with the following reagent concentrations: 1 $\times$  PCR buffer, 3.5 mM  $\text{MgCl}_2$ , 0.2 mM dNTPs, 0.4  $\mu\text{M}$  each primer,

1.25 U of *Taq* polymerase (5 U/ $\mu$ l), and 50 ng of template DNA. The DNA was amplified at the following conditions: initial denaturation at 94°C for 2 min; followed by 35 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 2 min, and extension at 72°C for 3 min; with a final extension of 72°C for 10 min. The final PCR product was diluted 1:30 and used as the template for the second PCR. All reagent concentrations and conditions remained the same as for the first reaction but primers R16F2n and R16R2 were used in place of primers P1 and TINT for the second PCR. The expected band size was 1,240 bp, which is too long for Illumina MiSeq technology. Thus, because general bacterial DNA can easily be amplified instead of phytoplasma DNA even while using phytoplasma-specific primers (Skrzeczowski et al. 2001), we decided to test the 52 PCR products generated above through Sanger Sequencing using the Biosystems SeqStudio Genetic Analyzer (Thermo Fisher Scientific) before proceeding with other sequencing options (i.e., Illumina High Throughput Sequencing using RNA).

To survey bacterial communities, primers 799F and 1199R, modified to contain Illumina overhang adapters, were used to amplify the V5 to V7 regions of the 16S rRNA because these primers have been shown to successfully avoid chloroplast and plant mitochondrial DNA after performing a gel extraction of the bacterial band (Horton et al. 2014). Based on a modified protocol from Beckers et al. (2016), the PCR assays were performed in 25  $\mu$ l, and each reaction contained 1 $\times$  PCR buffer with KCl, 2 mM MgCl<sub>2</sub>, 0.2 mM dNTP mix, 0.4  $\mu$ M each primer, 2 U of *Taq* polymerase (5 U/ $\mu$ l), and 10 ng of template DNA. The DNA samples were amplified under the following conditions: initial denaturation at 95°C for 3 min; followed by 35 cycles of denaturation at 95°C for 45 s, annealing at 54°C for 30 s, and extension at 72°C for 1 min; with a final extension performed at 72°C for 10 min. After PCR, products were separated on a 1.5% agarose gel in 1 $\times$  Tris-acetate-EDTA buffer stained with GelRed (VWR, Radnor, PA, U.S.A.). Two bands were produced at approximately 1,000 and 500 bp. To separate the mitochondria and chloroplast DNA (present at 1,000 bp) from bacterial DNA, the 500-bp band was excised from the gel, placed into 100  $\mu$ l of Milli-Q water, and heated to 90°C in the thermocycler for 5 min. A second PCR using the gel-purified product as template was carried out with the same conditions and reagent concentrations as the first PCR to purify only bacterial DNA.

To survey fungal communities, primers ITS1F and ITS2, modified to contain Illumina overhang adapters, were used to target ITS1 (Usyk et al. 2017). Each 25- $\mu$ l reaction contained 1 $\times$  PCR buffer, 3.5 mM MgCl<sub>2</sub>, 0.2 mM dNTP mix, 0.4  $\mu$ M each primer, 1.25 U of *Taq* polymerase, and 20 ng of template DNA (White et al. 1990). The DNA samples were run using the following conditions: initial denaturation at 95°C for 3 min; followed by 35 cycles of denaturation at 95°C for 30 s, annealing at 55°C for 30 s, and extension at 68°C for 2 min; with a final extension performed at 68°C for 10 min (Usyk et al. 2017). The expected band size was approximately 300 bp.

Separate aliquots of the DNA from each biological replicate were then used for analysis of the four libraries—that is, 16S (bacteria and phytoplasmata), 18S (nematodes), and ITS (fungi)—using the Nextera XT Index Kit (Illumina, San Diego, CA, U.S.A.), which was also used to add appropriate barcodes for multiplexing. Mock communities were generated for each library using the ZymoBIOMICS Microbial Community DNA Standard (Zymo Research, Irvine, CA, U.S.A.) and separate communities of known fungal DNA and known nematode DNA that were generated in house. Two negative controls were run with each library, one containing only nuclease-free water and the other containing an empty

DNA extraction to test for contamination in the extraction kits. The samples were sequenced on the Illumina MiSeq platform using a 2  $\times$  300 bp paired-end format yielding approximately 100,000 reads per sample.

**Bioinformatics processing.** The raw sequence data for each library type were processed separately in QIIME2 version 2019.7.0 (Bolyen et al. 2019). After merging all paired-end reads and trimming the adaptor and primer sequences, amplicon data read quality was assessed using the DADA2 pipeline, and low-quality bases, phiX reads, and chimeric reads were discarded. Taxonomy was assigned in QIIME2 using the SILVA v132 database at the 99% confidence level for the 16S and 18S libraries (Quast et al. 2013), while UNITE99 version 7 was used for the ITS library (Nilsson et al. 2019). Rarefaction curves were constructed for each sample individually from each library as well as per symptom type using QIIME2 (Supplementary Figs. S2 to S4). Samples were rarefied for each library by finding the median read frequency of all samples and choosing the number that also avoids losing the most samples.

All amplicon sequence variants (ASVs) corresponding to plant organelles or represented by fewer than 1,000 reads were removed prior to analyses. To characterize microbial communities associated with each of the different symptom types, both nonparametric and parametric tests were carried out using R version 3.6.2 (2019-12-12) (R Core Team 2019). The R package “phyloseq” (McMurdie and Holmes 2013) was used to create the framework to work with metabarcoding datasets and “ape” version 5.3 (Paradis and Schliep 2018) was used to interpret the phylogenetic relationships of the data, including computing the unweighted Unifrac distances for all ASVs. To test for taxa with differential abundance between symptoms and locations, we used the nonparametric Wilcoxon rank-sum test on the median read proportions for each taxon at all taxonomic ranks using the “agricolae” package (de Mendiburu 2020). We then used the package “metacoder” to parse, plot, and manipulate the taxonomic data sets to create differential heat trees as pairwise comparisons for visualization (Foster et al. 2017). Differential heat trees visualize and quantitatively depict important statistics associated with taxa in metabarcoding studies, including abundance through the use of color and size of nodes and branches in the taxonomic tree, which conventional graphing techniques cannot (Foster et al. 2017). Specifically, for our study, the package “DESeq2” (Love et al. 2014) was used to confirm the results from the heat tree analysis and specify the ASVs present in only symptomatic samples using the contrasts of naïve versus symptomatic samples. Both the “agricolae” and “DESeq2” programs used the Benjamini-Hochberg method to correct for false discovery rates in our analyses and provide an adjusted *P* value.

$\alpha$ -Diversity metrics were used to detect differences in the microbial communities of different symptom types, including Faith’s phylogenetic diversity using the package “PICANTE” (Kembel et al. 2010) and Shannon diversity and Simpson’s diversity, the latter using the package “vegan” in R (Oksanen et al. 2019). Within the same diversity index, differences between groups were tested using analysis of variance (ANOVA) at *P* < 0.05 and pairwise contrasts were calculated using the package “emmeans” (Lenth et al. 2020). Symptom type, location, and time all served as factors for the ANOVA to determine differences in the overall community composition. Nonmetric multidimensional scaling (NMDS) using unweighted Unifrac distance matrix was also calculated for the three different libraries using the “vegan” package in order to detect for significant clustering in  $\beta$ -diversity based on site or symptom type. Permutational multivariate analysis of variance (PERMANOVA) was used to identify significant differences in the  $\beta$ -diversity of communities between symptom type and location using 999 permutations and the unweighted Unifrac

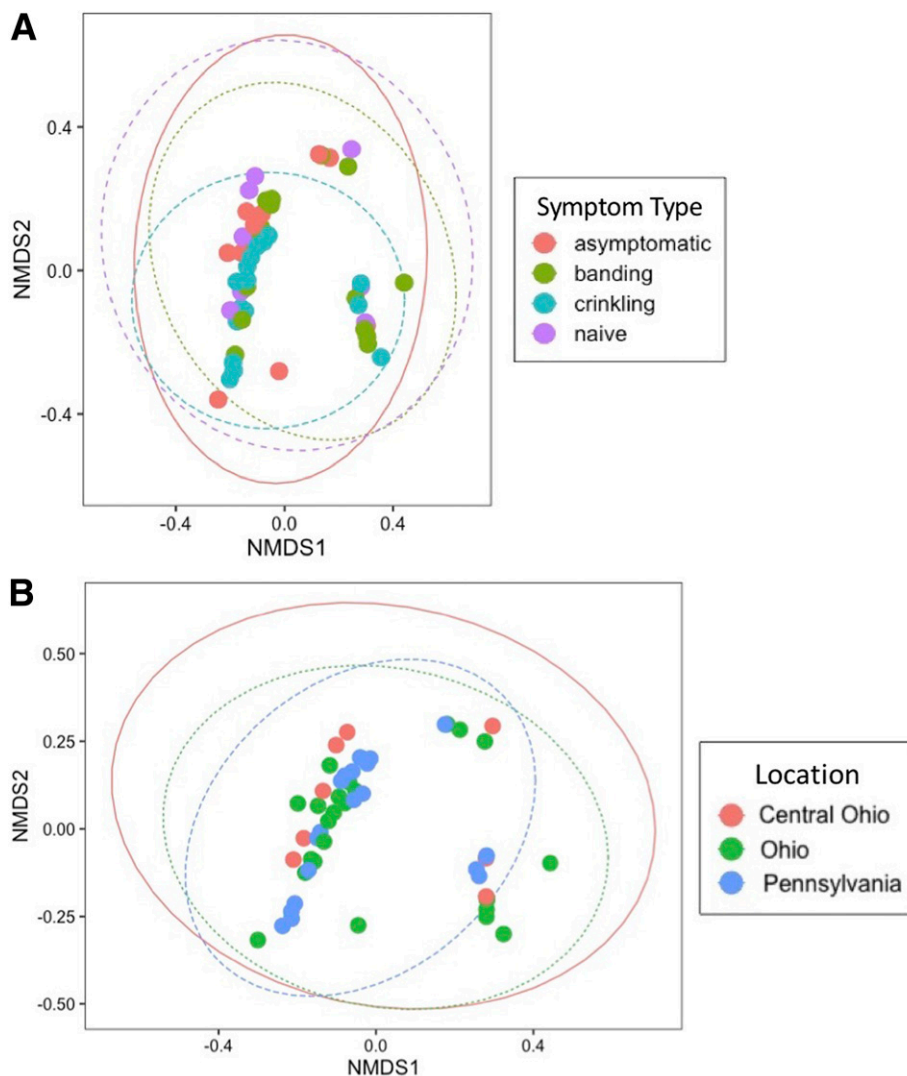
distance matrix. Symptom type and location were considered fixed effects in the model while time was used to stratify permutations.

**Indicator species analysis.** To seek further support for the results obtained through the pipeline described above, we also conducted an indicator species analysis (ISA) for each library type (Dufrene and Legendre 1997). Indicator species are considered the most characteristic species of a priori groups (i.e., those that are

found mostly in a single group or are present in the majority of sites belonging to that group) (Dufrene and Legendre 1997). Species are considered to be indicators of a specific site if they satisfy one of three requirements: (i) they reflect the biotic or abiotic state of the environment, (ii) they provide evidence for environmental change, or (iii) they influence the taxonomic diversity within an area (De Cáceres and Legendre 2009). ISA is frequently used by community ecologists but is encouraged for use by plant pathologists, particularly when evaluating potentially ambiguous causal agents and when investigating causal agents from emerging plant diseases (Severns and Sykes 2020). Analyses were carried out using the R package “indicspecies” version 1.7.8 (De Cáceres and Legendre 2009). ISA requires both a community data matrix and a classification of sites into specific a priori groups (De Cáceres and Legendre 2009). The indicator index is a product of two components labeled A and B, where component A is considered the positive predictor value and component B is considered the sensitivity value (Dufrene and Legendre 1997). Component A measures how likely the indicator species is to occur in only the designated group and component B measures how likely that species is to occur in all sites within that group (Dufrene and Legendre 1997). These values range from 0 to 1, with 1 meaning that the species only occurs

Source	df	SS	R <sup>2</sup>	F	P <sup>a</sup>
Symptom type	3	0.8965	0.10047	1.7871	0.025*
Plot location	2	0.4632	0.05191	1.4272	0.102
Symptom type × location	2	0.4240	0.04752	1.3066	0.212
Residuals	44	7.1396	0.80010	–	–
Total	51	8.9233	1	–	–

<sup>a</sup> Asterisk (\*) indicates significance measured at  $P < 0.05$ .



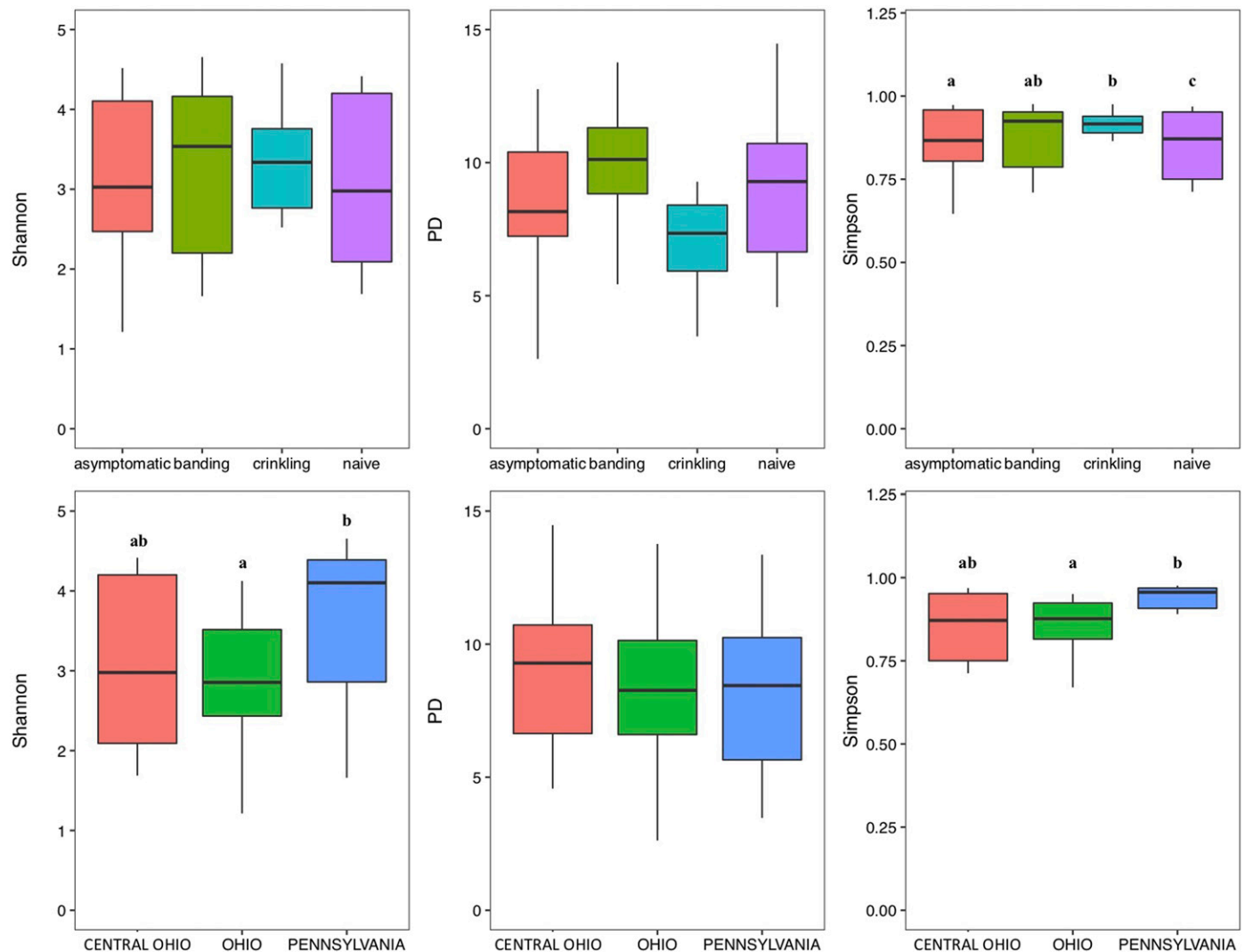
**Fig. 1.** Nonmetric multidimensional scaling (NMDS) charts do not show any obvious clustering in the bacterial community based on **A**, symptom type or **B**, location. Ellipses represent 95% confidence intervals.

in the designated groups (component A) and it occurs in all sites within that group (component B); thus, it is a true indicator species. Our samples were clustered based on symptom type for one analysis and were then reclustered based on location for the second analysis, and the ASVs of each library were analyzed independently. To further evaluate the statistical significance of the analysis, the sample types were also classified into four groups using nonhierarchical cluster analysis using  $k$  means, and the indicator species were determined from these groups. Because  $k$  means use centroids to define clusters that are most similar to one another, we used this algorithm to highlight whether another confounding factor could be acting on our data outside of the defined a priori groups. The significance of indicator species per group for all analyses was measured at  $P < 0.05$ .

## RESULTS

**Nematodes.** Sequences were trimmed to 240 bp to preserve the best-quality reads. After filtering out sequences identified as plant organelles, in total, 319,507 raw reads were retained. The samples were rarefied at 200 reads, reducing the total number of reads to 9,200. Because LCM is the only nematode reported to date in

American beech trees, and the eukaryotic microbiome of leaves is much less diverse than the fungal and bacterial microbiomes (Sivakumar et al. 2020), this low number of reads was expected. Upon clustering the DNA sequences into ASVs at the 99% confidence level, there were, in total, 231 ASVs, with 63 from the kingdom Metazoa and 68 from the kingdom Fungi. Out of the Metazoan ASVs, there were only three in the phylum Nematoda. After further filtering out those DNA sequences that matched with a less than 97% similarity, the most common putative ASV detected in the library corresponded to LCM, with 3,946 total reads. This ASV had a 100% identity to LCM identified by Carta and Li (2019) when run against the NCBI Basic Local Alignment Search Tool (BLAST) (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). When compared against the closest known type specimen, *Litylenchus crenatae* from Morioka, Iwate, Japan (Kanzaki et al. 2019), the ASV we detected for LCM had a sequence identity of 99.6%. It matched with the closest known nematode species, *L. coprosoma*, at 99.2% identity and the next closest genus of nematode, *Subanguina radicolica*, with 99.6% identity (Kanzaki et al. 2019). The ASV we detected for LCM was present in all symptom types with the following average read counts:  $129 \pm 32.46$  ( $n = 7$ ) reads for naïve,  $104.82 \pm 24.38$  ( $n = 11$ ) reads for asymptomatic,  $53.54$



**Fig. 2.** Diversity indices (Shannon, phylogenetic diversity [PD], and Simpson) for the 16S library based on symptom type and location. Significantly different means at  $P < 0.05$  are denoted by different letters.

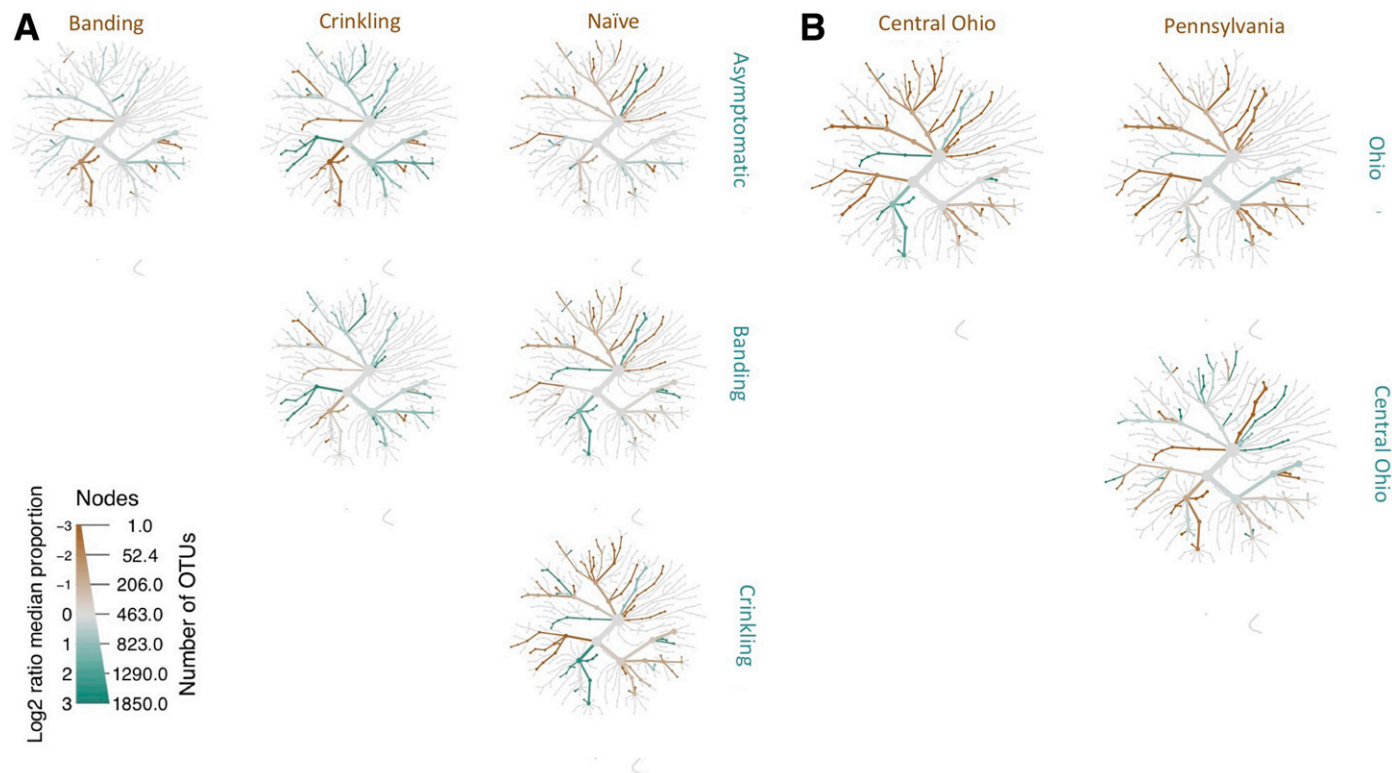
$\pm 21.66$  ( $n = 13$ ) reads for banding, and  $79.6 \pm 21.94$  ( $n = 15$ ) reads for crinkling.

Due to the small read numbers corresponding to other ASVs from the kingdom Metazoa, further analyses, including all diversity indices and generating NMDS plots, were not warranted. In addition, the differential abundance analysis indicated there were no significant ASVs that were more abundant in any symptom type.

The ISA, conducted with a priori groups defined by symptom type or location or when the samples were clustered based on  $k$  means, showed that no species was indicative of just one symptom type. However, the analysis showed that the ASV associated with LCM was present in all symptom types and locations.

**Bacteria and phytoplasmas.** Based on our results using Sanger Sequencing, only bacterial species not belonging to ‘*Candidatus* Phytoplasma’ were amplified when testing the phytoplasma-specific primers, which indicated that no phytoplasmas were detected in any of our samples. Thus, further screening for phytoplasmas using Illumina Sequencing was not conducted.

For the V5 to V7 16S rRNA regions specific for bacteria, sequences were trimmed using DADA2 at 240 bp to preserve sample quality; thus, only the V5 and V6 regions were analyzed. In total, 377,082 raw reads were retained. After rarefying the data at 4,380 reads/sample to retain the most samples, 226,544 reads remained and were clustered into 1,967 ASVs at the 99% confidence level. The most common putative ASV with 7,559 reads



**Fig. 3.** Differential heat trees show significant differences in the pairwise comparisons of bacterial amplicon sequence variants (ASVs) by **A**, symptom type and **B**, location based on the median read proportion. OTUs = operational taxonomic units. These graphical images allow us to visualize the differences in community abundance between sample types and determine the taxa that are over- or underrepresented, which could point toward a potential causal agent. Branches and nodes highlighted in blue represent that more ASVs of that genus were detected in the samples in the categories labeled on the right of the chart, and branches and nodes highlighted in brown represent that more ASVs were detected in the samples in the categories listed above the charts. Significance was determined using the Wilcoxon rank-sums test. The full taxonomic tree is provided in Supplementary Figure S4.

Top indicator species genera	<i>P</i> value <sup>a</sup>	Component A	Component B	Total read count
<i>Wolbachia</i> sp.	0.002**	1	1	5,297
<i>Wolbachia</i> sp.	0.014*	0.9996	1	7,559
<i>Pseudomonas</i> sp.	0.007**	0.9959	1	7,441
<i>Paenibacillus</i> sp.	0.008**	0.9944	1	1,186
<i>Erwinia</i> sp.	0.025*	0.9674	1	2,051
<i>Erwinia</i> sp.	0.05*	0.9947	0.875	1,511

<sup>a</sup> Asterisks \* and \*\* indicate significance measured at  $P < 0.05$  and  $0.01$ , respectively.

**TABLE 3**  
**Clusters used for the indicator species analysis based on *k* means (*k* = 4)**

Clusters, time <sup>a</sup>	Location	Symptoms
<b>Bacterial clusters</b>		
Cluster 1		
August	Northeast Ohio	Crinkling
Cluster 2		
May	Northeast Ohio	Crinkling
May	Pennsylvania	Crinkling
Cluster 3		
August	Northeast Ohio	Banding
August	Pennsylvania	Banding
August	Pennsylvania	Crinkling
Cluster 4		
May	Northeast Ohio	Asymptomatic
May	Northeast Ohio	Banding
May	Central Ohio	Naïve
May	Pennsylvania	Banding
August	Northeast Ohio	Asymptomatic
August	Central Ohio	Naïve
August	Pennsylvania	Asymptomatic
<b>Fungal clusters</b>		
Cluster 1		
August	Northeast Ohio	Crinkling
August	Northeast Ohio	Banding
Cluster 2		
May	Northeast Ohio	Asymptomatic
May	Central Ohio	Naïve
May	Pennsylvania	Banding
August	Central Ohio	Naïve
Cluster 3		
May	Northeast Ohio	Banding
May	Northeast Ohio	Crinkling
May	Pennsylvania	Crinkling
Cluster 4		
August	Northeast Ohio	Asymptomatic
August	Pennsylvania	Asymptomatic
August	Pennsylvania	Banding
August	Pennsylvania	Crinkling

<sup>a</sup> Sampling times were May or August.

corresponded to a *Wolbachia* sp. endosymbiont, while the second most common ASV, with 7,441 reads, was associated with the genus *Pseudomonas*.

The PERMANOVA determined that there were no significant groupings based on plot location or when symptom type and location were compared together; there were significant groupings based on symptom type alone, yet the  $R^2$  value was small, indicating only a slight effect on variation among bacterial communities (Table 1). The NMDS plots in the first two axes did not indicate any obvious clustering of ASVs when assigned symptom type or location (Fig. 1A and B). In terms of  $\alpha$ -diversity, there were no significant differences in Shannon diversity based on symptom type but there were significant differences based on location (Fig. 2). In regard to the Faith's phylogenetic diversity, the Kruskal-Wallis test showed that there were no significant differences in diversity based on symptom type or location (Fig. 2). Simpson's diversity analysis determined there was a significant difference in evenness based on symptom type and location (Fig. 2). Samples with crinkling symptoms had the greatest average evenness ( $D = 0.916 \pm 0.008$ ) and asymptomatic samples had the lowest average evenness index ( $D = 0.853 \pm 0.033$ ).

The differential abundance analysis indicated distinct differences in taxa among symptom types (Fig. 3A) and locations (Fig. 3B) based on the adjusted *P* values. Specifically, there were four ASVs that were present in only symptomatic samples (i.e., banding and crinkling symptom types). The four ASVs corresponded to the following genera: *Wolbachia*, *Pseudomonas*, *Erwinia*, and *Paenibacillus*. The full taxonomic tree for comparison is available in the supplementary material (Supplementary Fig. S5).

When the samples were assigned a priori groups based on symptom type for the ISA, six ASVs were identified as indicators: two *Wolbachia* spp., a *Pseudomonas* sp., a *Paenibacillus* sp., and two *Erwinia* spp. (Table 2). When a priori groups were defined by location, the ISA did not return any significant indicator species.

**TABLE 5**  
**Symptom type and location effects of fungal community composition of beech leaves as evidenced by permutational multivariate analysis of variance**

Source	df	SS	$R^2$	<i>F</i>	$P^a$
Symptom type	3	1.0188	0.15563	3.7201	0.001***
Plot location	2	1.0840	0.16559	5.9235	0.001***
Symptom type × location	2	0.4175	0.06378	2.2817	0.001***
Residual	44	4.0260	0.61500	–	–
Total	51	6.5463	1	–	–

<sup>a</sup> Asterisks (\*\*\*) indicate significance measured at  $P < 0.001$ .

**TABLE 4**  
**Indicator species for the bacterial microbiome defined by *k* means**

Top indicator species genera	<i>P</i> value <sup>a</sup>	Component A	Component B	Total read count
<b>Clusters 1, 2, and 3</b>				
<i>Wolbachia</i> sp.	0.001***	0.992	1	5,297
<i>Wolbachia</i> sp.	0.013*	0.9983	1	7,559
<i>Pseudomonas</i> sp.	0.008**	0.9645	1	7,441
Uncultured endophytic bacterium	0.048*	0.939	1	1,796
<b>Clusters 1, 3, and 4</b>				
Rhizobiales	0.011*	1	1	1,168

<sup>a</sup> Asterisks \*, \*\*, and \*\*\* indicate significance measured at  $P < 0.05$ , 0.01, and 0.001, respectively.

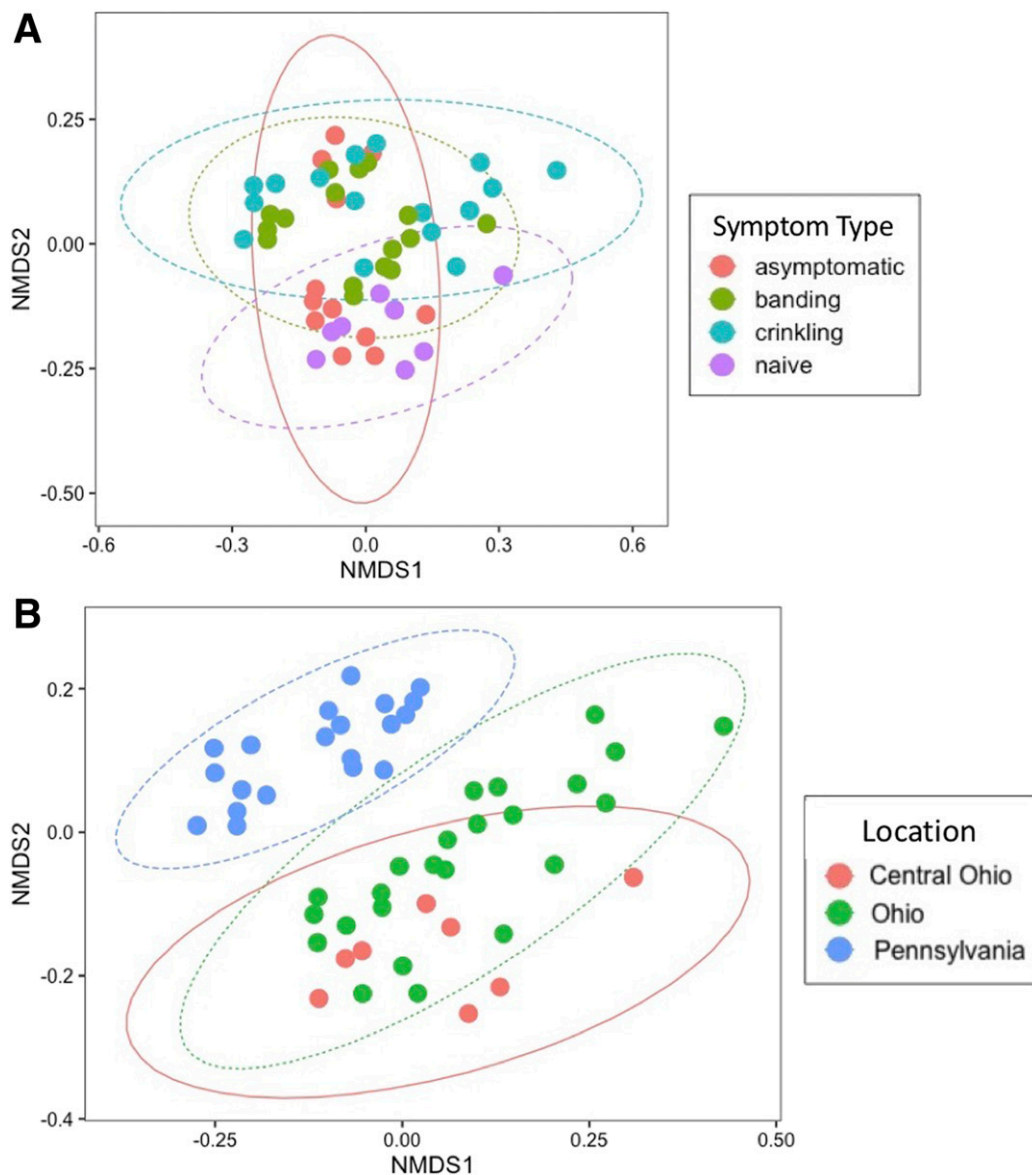
Based on  $k$  means, four clusters were created for the ISA of bacteria (Table 3). No species were associated with just one cluster but several species were indicators when multiple clusters were analyzed together (Table 4). The indicator species for clusters one, two, and three together were two ASVs corresponding to *Wolbachia* spp., one associated with a *Pseudomonas* spp., and one associated with an uncultured endophytic bacterium (Table 4). The indicator species for clusters one, three, and four, when analyzed together, were predicted to be from the order Rhizobiales.

**Fungi.** Samples were trimmed at 250 bp to preserve sample quality and, in total, 3,161,686 raw reads were obtained. The data were rarefied at 23,500 reads/sample to preserve the most samples; thus, 1,222,000 reads were retained and clustered into 6,273 ASVs, with the majority of ASVs corresponding to ascomycetous fungi (5,376 ASVs) based on 99% confidence.

The PERMANOVA showed that there were significant differences in putative ASVs among both symptom types and plot locations (Table 5). The NMDS analysis showed strong clustering of

ASVs based on location, specifically Ohio versus Pennsylvania, but not symptom type based on the ordination of the first two axes (Fig. 4A and B). The Shannon and Faith's phylogenetic diversity analyses showed that there were significant differences in diversity among both symptom types and locations (Fig. 5). The Simpson diversity determined that the samples differed in community evenness based on symptom type but not location (Fig. 5). The average evenness index was greatest for asymptomatic samples ( $D = 0.944 \pm 0.005$ ) and lowest for crinkling samples ( $D = 0.894 \pm 0.018$ ).

According to the Wilcoxon rank-sum test, there were slight differences in the taxa abundance among symptom types (Fig. 6A) and locations (Fig. 6B), specifically between the Ohio sites and Pennsylvania sites, based on the adjusted  $P$  values. However, there was only one ASV, corresponding to the genus *Paraphaeosphaeria*, that was present in only symptomatic and not asymptomatic material. The full taxonomic tree can be viewed in the supplementary material (Supplementary Fig. S6).



**Fig. 4.** Nonmetric multidimensional scaling (NMDS) charts do not show obvious clustering in fungal community samples based on **A**, symptom type, but clustering is noted based on **B**, location. Ellipses represent 95% confidence intervals.

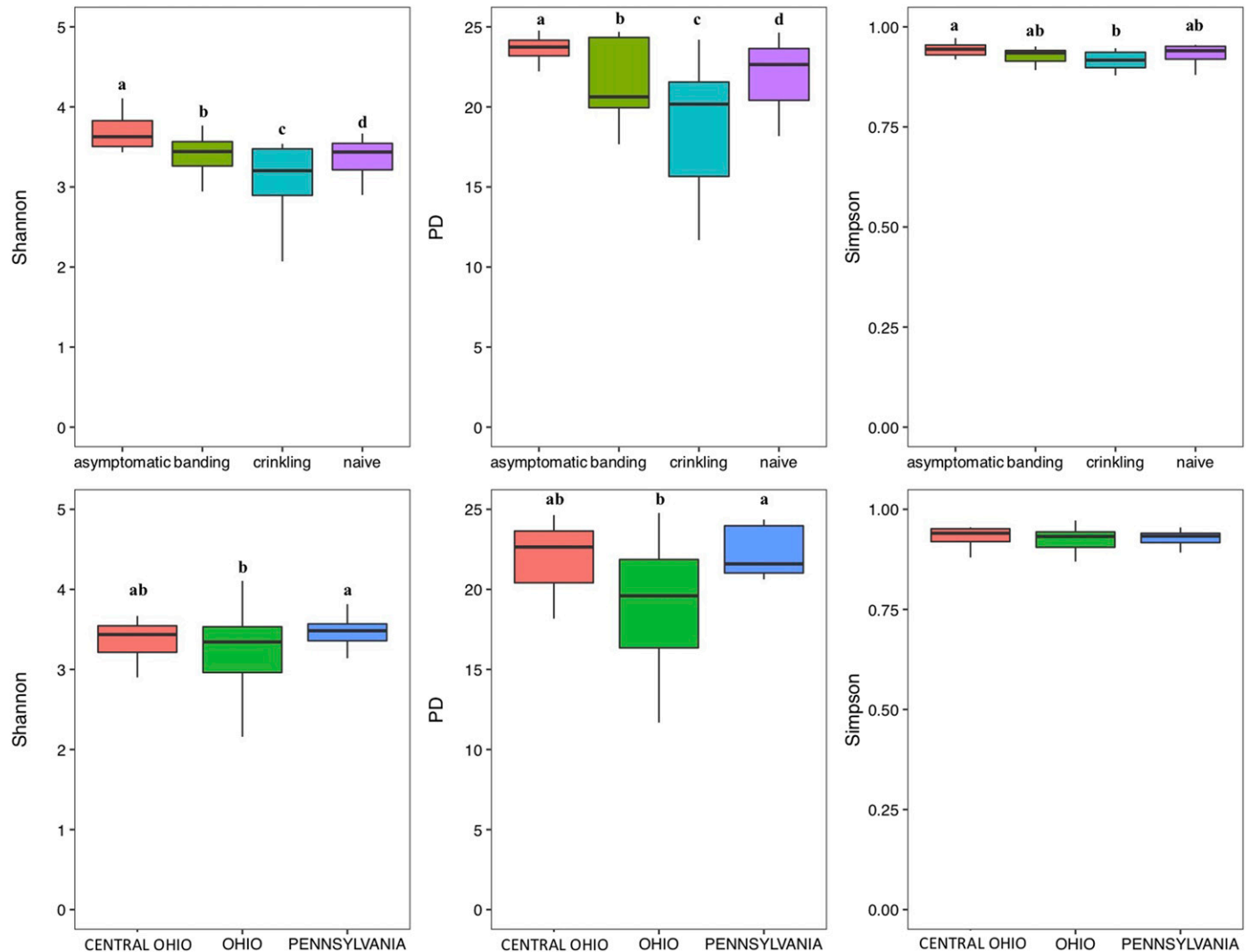


The ISA results for the symptom groups are shown in Table 6. Using a priori groups based on symptom type revealed that the ASV associated with an uncultured fungus was the indicator species for asymptomatic samples. There were four ASVs associated with the naïve symptom type corresponding to *Dothiora* sp., *Phoma* sp., and two uncultured fungi. One ASV linked to the genus *Chaetothyriales* was significantly associated with combined asymptomatic and naïve samples. There were no indicator species detected for banding or crinkling symptoms independently but two ASVs corresponding to *Paraphaeosphaeria* sp. and *Diaporthe* sp. were indicative of banding and crinkling samples when analyzed together. When the asymptomatic, banding, and crinkling samples were analyzed together, 10 ASVs were determined to be indicator species, including three uncultured fungi, three ASVs corresponding to *Ramularia* sp., a *Geastrumia* sp., a species in the subclass Lecanoromycetidae, an *Orbiliaceae* sp., and a *Coniozyma* sp. Finally, one ASV associated with *Paraconiothyrium* sp. was determined to be an indicator species for the asymptomatic, crinkling, and naïve samples when analyzed together.

The ISA results for the location groups are shown in Table 7. Five ASVs were indicative of samples from central Ohio. The

genera associated with these ASVs were *Dothiora* and *Didymella*, plus three uncultured fungi. The indicator species of the Pennsylvania samples were three ASVs corresponding to *Ramularia* spp., one associated with *Curvibasidium* sp., and seven ASVs identified as uncultured fungi. Three ASVs identified as uncultured fungi were determined to be indicator species of samples from northeastern Ohio. Ten ASVs were indicator species of the northeast Ohio and central Ohio sites, when analyzed together. The ASVs were identified as species within the following genera: *Plagiostoma*, *Paraconiothyrium*, *Diaporthe*, *Fusicoccum*, *Seimatosporium*, and *Toxicocladosporium*, plus three uncultured fungal species. One ASV corresponding to an uncultured fungus was determined to be an indicator species of central Ohio and Pennsylvania sites when analyzed together. When the northeast Ohio and Pennsylvania sites were analyzed together, three ASVs associated with the genus *Ramularia*, one ASV in the subclass Lecanoromycetidae, one ASV associated with the family Orbiliaceae, one ASV in the genus *Geastrumia*, one ASV in the genus *Coniozyma*, plus three ASVs associated with uncultured fungi were all considered indicators.

Four groups were clustered based on the *k* means (Table 3), and the ISA indicated that there was only one species that corresponded



**Fig. 5.** Diversity indices (Shannon, phylogenetic diversity [PD], and Simpson) for the internal transcribed spacer library based on symptom type and location. Significantly different means at  $P < 0.05$  are denoted by different letters.

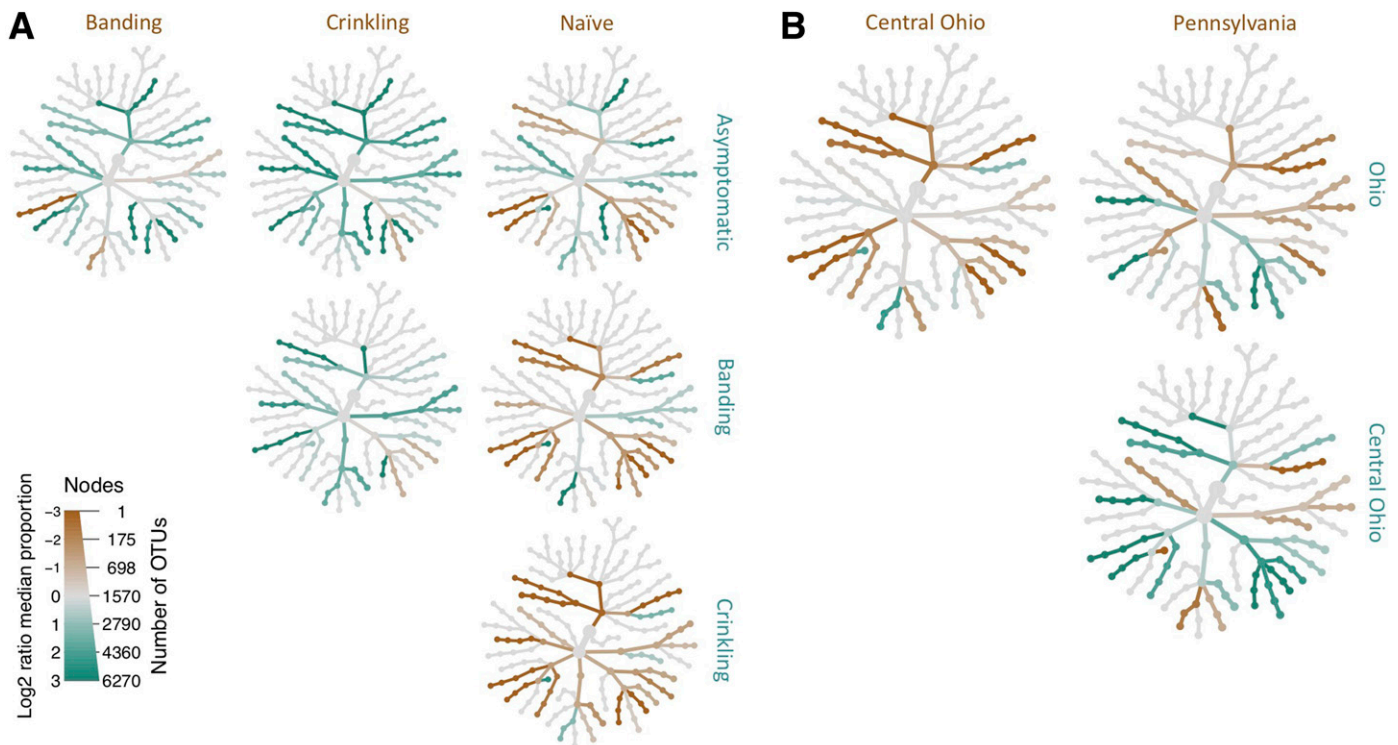
to just one cluster, four species that corresponded to two clusters together, and five species that corresponded to three clusters (Table 8). The significant indicator species for cluster four was an ASV associated with an uncultured fungus. One ASV putatively identified as a species in the genus *Paraphaeosphaeria* was determined to be an indicator species of clusters one and three when analyzed together. Three ASVs were determined to be significant indicators for clusters two and three and three when analyzed together, which corresponded to the genera *Megacollybia* and *Gymnopus* and an uncultured fungus. When clusters two and four were analyzed together, one ASV associated with the genus *Chaetothyriales* was the indicator species. Two ASVs identified as uncultured fungi and one in the genus *Seimatosporium* were identified as indicator species for clusters one, two, and three when analyzed together. When comparing clusters one, two, and four, the indicator species was identified as an uncultured fungus, as was the indicator species for clusters two, three, and four.

## DISCUSSION

The presence of LCM DNA in both symptomatic and asymptomatic foliage reported here is a confirmation of findings of live nematodes in both symptomatic and asymptomatic foliage by Reed et al. (2020) and molecular amplification by Burke et al. (2020). This suggests that the suspected nematode causal agent, LCM, may not be fully responsible for BLD. Reed et al. (2020) also speculate in their study that full BLD symptom development may need an additional disease agent along with LCM, which is consistent with our study results. The authors also mention that symptom development could require infection of buds or immature leaves by

LCM (Reed et al. 2020). In addition, LCM was not considered an indicator species for any of the clusters generated for the ISA based on the *k* means or when the groupings were defined by symptom type or location, yet it did indicate that the ASV corresponding to LCM was the most prevalent ASV associated with all sample types.

Based on the bacterial analyses, four ASVs associated with the genera *Wolbachia*, *Pseudomonas*, *Erwinia*, and *Paenibacillus* were detected in only the symptomatic material through both the differential abundance analysis and the ISA, suggesting that those ASVs may be associated with BLD in some way. *Wolbachia* endosymbionts belonging to the order Rickettsiales are obligate intracellular bacteria that are most often present in arthropods and filarial nematodes (Baldo et al. 2002). This relationship is typically mutualistic and almost always obligate between filarial nematodes and *Wolbachia* endosymbionts (Baldo et al. 2002). The *Wolbachia* spp. increase the fitness of the nematode by assisting with development, reproduction, and overall survival (Slatko et al. 2014). In 2009, a *Wolbachia* sp. was detected in the plant-parasitic nematode genus *Radopholus* and in 2016 in the genus *Pratylenchus* (Wasala et al. 2019). Brown et al. (2016) suggest that the *Wolbachia* sp. found in the *Pratylenchus* nematode increases the fitness of the nematode through nutrient supplementation, specifically for heme, making it more pathogenic. Although it has not been proven, a similar mutualism could be occurring between LCM and the *Wolbachia* spp. identified in infected beech material that could be contributing to BLD, and its presence in symptomatic tissue could explain the differences in symptomatology in the leaves. *Wolbachia* is also one of the three genera of parthenogenesis-inducing bacteria, which occurs when bacteria manipulate haplodiploid sex determination of their host by converting haploid



**Fig. 6.** Differential heat trees show significant differences in the pairwise comparisons of fungal amplicon sequence variants (ASVs) by **A**, symptom type and **B**, location based on the median read proportion. OTUs = operational taxonomic units. Branches and nodes highlighted in blue represent that more ASVs of that genus were detected in the samples in the categories labeled on the right of the chart, and branches and nodes highlighted in brown represent that more ASVs were detected in the samples in the categories listed above the charts. Significance was determined using the Wilcoxon rank-sums test. The full taxonomic tree is provided in Supplementary Figure S5.

cells into female diploid cells (Brucker and Bordenstein 2012). Species within the genus are also known to induce cytoplasmic incompatibility in its insect host, which typically results in postmating reproductive isolation through F<sub>1</sub> inviability in crosses between infected males and uninfected females (Brucker and Bordenstein 2012). These effects lead to isolation of populations and increased host speciation (Brucker and Bordenstein 2012); therefore, we speculate that the *Wolbachia* sp. associated with BLD could be indicating a newly emerged population or subspecies of *L. crenatae*.

As previously mentioned, a similar study by Burke et al. (2020) determined the presence of *Wolbachia* in only symptomatic material. Upon further investigation, 14 operational taxonomic units from Burke et al. (2020) identified as putative species of *Wolbachia* matched the putative species identified in our study. Although we speculate that *Wolbachia* could be in a symbiotic relationship with LCM, Burke et al. (2020) speculated that mites or ambrosia beetles that are also common hosts of *Wolbachia* could possibly be transporting LCM eggs between tree hosts, which would explain the presence of the bacteria on symptomatic leaves. In contrast to our study, they also detected the bacterial genera *Methylocystis*, *Mucilaginibacter*, and *Terimonas* in only symptomatic material (Burke et al. 2020). Although interesting, these bacterial genera have never been reported as phytopathogenic; thus, they likely do not contribute to BLD development.

In addition to *Wolbachia*, the genera *Pseudomonas* and *Erwinia* that were also identified in our differential abundance analysis are

commonly found as plant pathogens; therefore, it is possible that one of these species could be causally associated with BLD (Mansfield et al. 2012). Other research has shown that species in both of these genera have a relationship with nematodes. For example, there is evidence that the pathogen *Erwinia amylovora* (Burrill 1882) Winslow, emend. Hauben et al., which causes fire blight of pear and apple trees, can live inside the digestive tract of saprozoic nematodes and be transported to susceptible host species to cause disease (Chantanao and Jensen 1969). In addition, the ubiquitous and potentially opportunistic *Pseudomonas putida* (Trevisan 1889) Mígula 1895 (Fernández et al. 2015) and *P. fluorescens* Mígula 1895 have been shown to have a symbiotic relationship with the pine wood nematode species *Bursaphelenchus xylophilus* (Steiner & Bührer) Nickle, which causes pine wilt (Fu et al. 2020; Zhao and Lin 2005). Fu et al. (2020) showed that these bacterial species increase the nematode's reproduction and, thus, increase the overall virulence of *B. xylophilus*. All three of these bacterial species (*E. amylovora*, *P. putida*, and *P. fluorescens*) were among the top BLASTn matches for the two ASVs identified in our research; therefore, additional research on their relationship with LCM should be pursued.

Interestingly, we did not see significant differences in bacterial communities between samples for the Faith's phylogenetic or Shannon diversity indices based on symptom type but we did for the Simpson's diversity index. This indicates that the evenness of the bacterial communities differs among symptom type; thus, it is possible that symptomatic leaves have an overabundance of a

**TABLE 6**  
Indicator species for the mycobiome using a priori groups based on symptom type

Symptom, top indicator species genera	P value <sup>a</sup>	Component A	Component B	Total read count
Naïve				
Uncultured fungus	0.018*	0.9993	1	1,959
Uncultured fungus	0.016*	0.9993	1	1,461
<i>Dothiora</i>	0.017*	0.9882	1	1,140
<i>Phoma</i>	0.022*	0.863	1	2,038
Asymptomatic				
Uncultured fungus	0.034*	0.9134	1	2,318
Naïve and asymptomatic				
Chaetothyriales (order)	0.032*	0.9977	1	5,862
Banding and crinkling				
<i>Paraphaeosphaeria</i>	0.002**	1	1	2,233
<i>Diaporthe</i>	0.014*	0.9893	1	1,492
Asymptomatic, banding, and crinkling				
<i>Ramularia</i>	0.015*	1	1	62,291
<i>Ramularia</i>	0.015*	1	1	48,883
Uncultured fungus	0.015*	1	1	18,844
<i>Ramularia</i>	0.015*	1	1	8,897
Uncultured fungus	0.015*	1	1	1,506
Uncultured fungus	0.002**	0.9992	1	6,781
Lecanoromycetidae (subclass)	0.008**	0.9981	1	10,042
<i>Orbiliaceae</i>	0.029*	0.9884	1	1,786
<i>Geastrumia</i>	0.016*	0.9875	1	1,077
<i>Coniozoma</i>	0.047*	0.9605	1	2,886
Asymptomatic, naïve, and crinkling				
<i>Paraconiothyrium</i>	0.022*	0.9507	1	3,499

<sup>a</sup> Asterisks \* and \*\* indicate significance measured at  $P < 0.05$  and  $0.01$ , respectively.

**TABLE 7**  
**Indicator species for the mycobiome using a priori groups based on location**

Location, top indicator species genera <sup>a</sup>	<i>P</i> value <sup>b</sup>	Component A	Component B	Total read count
Central Ohio				
Uncultured fungus	0.012*	0.9997	1	1,959
Uncultured fungus	0.015*	0.9995	1	1,461
<i>Dothiora</i> sp.	0.009**	0.9941	1	1,140
Uncultured fungus	0.029*	0.9681	1	6,377
<i>Didymella</i> sp.	0.008**	0.9265	1	2,038
NE Ohio				
Uncultured fungus	0.001***	1	1	1,454
Uncultured fungus	0.033*	0.8571	1	1,136
Uncultured fungus	0.049*	0.9132	0.8333	3,518
NW Pennsylvania				
<i>Ramularia</i> sp.	0.002**	1	1	14,699
<i>Ramularia</i> sp.	0.003**	0.9994	1	9,920
Uncultured fungus	0.005**	0.9924	1	6,660
<i>Ramularia</i> sp.	0.003**	0.9881	1	2,712
Uncultured fungus	0.013*	0.9733	1	1,787
Uncultured fungus	0.012*	0.9713	1	1,989
<i>Curvibasidium</i> sp.	0.032*	0.9562	1	2,457
Uncultured fungus	0.002**	0.9543	1	1,331
Uncultured fungus	0.008**	0.9216	1	2,407
Uncultured fungus	0.015*	0.8878	1	1,656
Uncultured fungus	0.043*	1	0.8	1,764
Central Ohio and NE Ohio				
<i>Plagiostoma</i> sp.	0.002**	1	1	2,815
<i>Paraconiothyrium</i> sp.	0.002**	1	1	1,007
<i>Fusicoccum</i> sp.	0.002**	0.9992	1	7,951
<i>Diaporthe</i> sp.	0.001***	0.992	1	1,057
Uncultured fungus	0.011*	0.9942	1	9,274
<i>Seimatosporium</i> sp.	0.004**	0.9889	1	3,867
<i>Toxicocladosporium</i> sp.	0.002**	0.9854	1	1,905
Uncultured fungus	0.027*	0.9548	1	4,473
Uncultured fungus	0.035*	1	0.875	1,799
Uncultured fungus	0.012*	1	0.875	1,475
Central Ohio and NW Pennsylvania				
Uncultured fungus	0.01**	0.9237	1	1,375
NE Ohio and NW Pennsylvania				
<i>Ramularia</i> sp.	0.015*	1	1	62,291
<i>Ramularia</i> sp.	0.015*	1	1	48,883
Uncultured fungus	0.015*	1	1	18,844
<i>Ramularia</i> sp.	0.015*	1	1	8,897
Uncultured fungus	0.015*	1	1	1,506
Uncultured fungus	0.002**	0.9988	1	6,781
Lecanoromycetidae (subclass)	0.001***	0.9975	1	10,042
<i>Orbiliaceae</i> sp.	0.014**	0.9839	1	1,786
<i>Geastrumia</i> sp.	0.022*	0.9808	1	1,077
<i>Coniozyma</i> sp.	0.019*	0.9456	1	2,886

<sup>a</sup> NE = northeast and NW = northwest.

<sup>b</sup> Asterisks \*, \*\*, and \*\*\* indicate significance measured at  $P < 0.05$ , 0.01, and 0.001, respectively.

pathogenic microbe that could be contributing to BLD. Alternatively, beneficial microbes may be suppressed on diseased trees so that they are more susceptible to infection. However, there were significant differences in diversity in the Shannon and Simpson's diversity indices among sample types based on location. When looking at the pairwise comparisons of the Simpson and Shannon diversity, only northwest Pennsylvania and northeast Ohio differed significantly, which we would expect, because these communities are located in two different geographical areas and forest settings (i.e., protected game land versus managed metropark); thus, the microbiome is likely to be different (Schlegel et al. 2018).

The putative fungal genus *Paraphaeosphaeria* that is significantly associated with only symptomatic samples includes species of endophytic fungi with cytotoxic compounds, plant pathogens, and mycoparasites (Baroncelli et al. 2020). Currently, there are no known cases of a pathogen from this genus affecting tree species but *Paraphaeosphaeria* spp. are found on a wide range of dicotyledonous and monocotyledonous plants and cause a variety of symptoms on leaves, including leaf spot, necrosis, and lesions (Lee et al. 2005). Regardless, both the differential abundance analysis and the ISA indicated that the ASV corresponding to the *Paraphaeosphaeria* genus was significantly associated with only samples showing symptoms of BLD, suggesting that it could be associated with the disease. However, because *Paraphaeosphaeria* spp. utilize chitinases during mycoparasitic modes (Giczey et al. 2001), it is possible that the fungus could simply be exploiting chitinous LCM eggs as a hyperparasite but have no actual role in BLD etiology (Mercer et al. 1992).

It is clear that there are significant differences in the fungal communities based on the Shannon, Faith's phylogenetic, and Simpson's evenness diversity indices among both symptom types and location of samples. The pairwise comparisons of both the evenness and phylogenetic diversities produced similar results, with asymptomatic and symptomatic samples differing significantly. Clear clustering of samples in the ITS NMDS plots was evident based on location, as would be expected based on the different

geographical areas and microclimates where they are found. There were also several ASVs that were identified as indicator species for specific locations, unlike the bacterial ISA, where no species were identified as significant based on location type. Three of the same fungal ASVs were associated with both central Ohio and the naïve symptom type, which further supports the suggestions that the mycobiome seems to be more closely associated with location than symptom type, because naïve samples were only collected in central Ohio. However, the genus *Paraphaeosphaeria* was an indicator for only banding and crinkling symptoms in the ISA, which supports the results of the differential abundance analysis and leads to speculation about whether or not it could be the causal agent of BLD.

Although our results show that the five aforementioned bacterial and fungal genera are associated only with symptomatic tissues, suggesting that they may be among the causal agents, we cannot ignore the alternative hypothesis (i.e., that these microorganisms are a consequence, and not the cause, of the disease). Indeed, with the exception of *Wolbachia*, these genera are known to harbor saprophytic species (McSpadden Gardener 2004; Preston 2004; Taylor and Crous 2001; Zhao and Qi 2011). In other words, it is possible that the putative species identified by the ASVs detected in our sequencing results, including LCM, are opportunistic pathogens or secondary colonizers of the dying symptomatic leaf tissues. However, this possibility does not explain the presence of the ASV corresponding to genus *Wolbachia*, given that it is an endosymbiont in nature and would not colonize as a saprophyte. This alternative hypothesis also seems unlikely given the potential pathogenic nature of many species within these genera and their ability to form relationships with nematodes. This uncertainty may be clarified by future pathogenicity studies using isolates of these microorganisms.

**Conclusion.** The findings of the differential abundance analysis suggest that multiple microbial players may be involved in development of BLD. Further investigations into the bacterial and fungal species present in only symptomatic material are needed to

**TABLE 8**  
Indicator species for the mycobiome defined by *k* means

Cluster, predicted top indicator species	<i>P</i> value <sup>a</sup>	Component A	Component B	Total read count
Cluster 4				
Uncultured fungus	0.008**	0.9694	1	2,318
Clusters 1 and 3				
<i>Paraphaeosphaeria</i> sp.	0.03*	0.8724	1	2,233
Clusters 2 and 3				
Uncultured fungus	0.002**	0.9968	1	1,188
<i>Megacollybia</i> sp.	0.047*	1	0.8571	1,485
<i>Gymnopus</i> sp.	0.047*	0.992	0.8571	2,231
Clusters 2 and 4				
<i>Chaetothyriales</i> sp.	0.028*	0.9983	1	5,862
Clusters 1, 2, and 3				
Uncultured fungus	0.049*	0.9961	1	9,274
Uncultured fungus	0.016*	0.98	1	2,488
<i>Seimatosporium</i> sp.	0.041*	0.9436	1	3,867
Clusters 1, 2, and 4				
Uncultured fungus	0.014*	0.9934	1	1,337
Clusters 2, 3, and 4				
Uncultured fungus	0.023*	0.9935	1	1,506

<sup>a</sup> Asterisks \*, \*\*, and \*\*\* indicate significance measured at *P* < 0.05, 0.01, and 0.001, respectively.

determine whether one or all of these species can cause symptoms of BLD, or if there is an association with LCM. This would entail isolating the microorganisms and attempting to prove pathogenicity through Koch's postulates. In addition, investigating the internal microbiome of LCM is necessary in order to determine which, if any, microbial species are able to live symbiotically inside the nematode. In sum, our results are consistent with prior hypotheses that LCM may be a necessary yet not sufficient component for BLD disease development.

**Data accessibility.** DNA sequences can be found under BioProject ID PRJNA665708 (<https://www.ncbi.nlm.nih.gov/bio/project/PRJNA665708>) and sequence processing and statistical analysis pipelines can be found at [https://github.com/CFearer/BLD\\_Microbial\\_Profiling](https://github.com/CFearer/BLD_Microbial_Profiling).

## ACKNOWLEDGMENTS

Sequencing was performed by the Ohio Agricultural Research and Development Center Molecular and Cellular Imaging Center.

## LITERATURE CITED

- Baldo, L., Bartos, J. D., Werren, J. H., Bazzocchi, C., Casiraghi, M., and Panelli, S. 2002. Different rates of nucleotide substitutions in *Wolbachia* endosymbionts of arthropods and nematodes: Arms race or host shifts? *Parassitologia* 44:179-187.
- Baroncelli, R., Da Lio, D., Vannacci, G., and Sarrocco, S. 2020. Genome resources for the endophytic fungus *Paraphaeosphaeria sporulosa*. *Mol. Plant-Microbe Interact.* 33:1098-1099.
- Becker, B., Op De Beeck, M., Thijs, S., Truyens, S., Weyens, N., Boerjan, W., and Vangronsveld, J. 2016. Performance of 16S rDNA primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies. *Front. Microbiol.* 7:650.
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., et al. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37:852-857.
- Brown, A. M. V., Wasala, S. K., Howe, D. K., Peetz, A. B., Zasada, I. A., and Denver, D. R. 2016. Genomic evidence for plant-parasitic nematodes as the earliest *Wolbachia* hosts. *Sci. Rep.* 6:34955.
- Brucker, R. M., and Bordenstein, S. R. 2012. Speciation by symbiosis. *Trends Ecol. Evol.* 27:443-451.
- Burke, D. J., Hoke, A. J., and Koch, J. 2020. The emergence of beech leaf disease in Ohio: Probing the plant microbiome in search of the cause. *For. Pathol.* 50:e12579.
- Carta, L. K., Handoo, Z. A., Li, S., Kantor, M., Bauchan, G., McCann, D., Gabriel, C. L., Yu, Q., Reed, S., Koch, J., Martin, D., and Burke, D. J. 2020. Beech leaf disease symptoms caused by newly recognized nematode subspecies *Litylenchus crenatae mccannii* (Anguinata) described from *Fagus grandifolia* in North America. *For. Pathol.* 50:e12580.
- Carta, L. K., and Li, S. 2019. PCR amplification of a long rDNA segment with one primer pair in agriculturally important nematodes. *J. Nematol.* 51:1-8.
- Chantanao, A., and Jensen, H. J. 1969. Saprophytic nematodes as carriers and disseminators of plant pathogenic bacteria. *J. Nematol.* 1:216-218.
- De Cáceres, M., and Legendre, P. 2009. Associations between species and groups of sites: Indices and statistical inference. *Ecology* 90:3566-3574.
- de Mendiburu, F. 2020. *agricolae*: Statistical procedures for agricultural research. R package version 1.3-2.
- Dufrene, M., and Legendre, P. 1997. Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecol. Monogr.* 67:345-366.
- Ewing, C. J., Hausman, C. E., Pogacnik, J., Slot, J., and Bonello, P. 2019. Beech leaf disease: An emerging forest epidemic. *For. Pathol.* 49:e12488.
- Fernández, M., Porcel, M., de la Torre, J., Molina-Henares, M. A., Daddaoua, A., Llamas, M. A., Roca, A., Carriel, V., Garzón, I., Ramos, J. L., Alaminos, M., and Duque, E. 2015. Analysis of the pathogenic potential of nosocomial *Pseudomonas putida* strains. *Front. Microbiol.* 6:871.
- Foster, Z., Sharpton, T., and Grunwald, N. 2017. Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. *PLOS* 13:e1005404.
- Fu, Y. M., Liu, H. B., and Wu, X. Q. 2020. Diversity and function of endobacteria in *Bursaphelenchus xylophilus* from *Pinus massoniana* Lamb. in different regions. *Forests* 11:487.
- Giczey, G., Kerényi, Z., Fülöp, L., and Hornok, L. 2001. Expression of cmg1, an exo- $\beta$ -1, 3-glucanase gene from *Coniothyrium minitans*, increases during sclerotial parasitism. *Appl. Environ. Microbiol.* 67:865-871.
- Gortari, M. C., and Hours, R. A. 2008. Fungal chitinases and their biological role in the antagonism onto nematode eggs. A review. *Mycol. Prog.* 7:221-238.
- Gundersen, D. E., and Lee, I.-M. 1996. Ultrasensitive detection of phytoplasmas by nested-PCR assays using two universal primer pairs. *Phytopathol. Mediterr.* 35:144-151.
- Horton, M. W., Bodenhausen, N., Beilsmith, K., Meng, D., Muegge, B. D., Subramanian, S., Vetter, M. M., Vilhjálmsson, B. J., Nordborg, M., Gordon, J. I., and Bergelson, J. 2014. Genome-wide association study of *Arabidopsis thaliana* leaf microbial community. *Nat. Commun.* 5:5320.
- Kanzaki, N., Ichihara, Y., Aikawa, T., Ekino, T., and Masuya, H. 2019. *Litylenchus crenatae* n. sp. (Tylenchomorpha: Anguinidae) a leaf gall nematode parasitising *Fagus crenata* Blume. *Nematology* 21:5-22.
- Kembel, S. W., Cowan, P. D., Helmus, M. R., Cornwell, W. K., Morlon, H., Ackerly, D. D., Blomber, S. P., and Webb, C. O. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26:1463-1464.
- Lee, H. B., Kim, K. M., and Jung, H. S. 2005. *Paraphaeosphaeria recurvifoliae*, a new species causing leaf spots and necrosis on *Yucca recurvifolia*. *Fungal Divers.* 20:71-81.
- Lenth, R., Buerkner, P., Herve, M., Love, J., Riebl, H., and Singmann, H. 2020. emmeans: Estimated marginal means, aka least-squares means. R package version 1.4.4. <https://cran.r-project.org/web/packages/emmeans/index.html>
- Love, M. I., Huber, W., and Anders, S. 2014. Moderation estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550.
- Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V., Machado, M. A., Toth, I., Salmond, G., and Foster, G. D. 2012. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol. Plant Pathol.* 13:614-629.
- Marra, R. E., and LaMondia, J. A. 2020. First report of beech leaf disease, caused by the foliar nematode *Litylenchus crenatae mccannii*, on American beech (*Fagus grandifolia*) in Connecticut. *Plant Dis.* 104:2527.
- McMurdie, P. J., and Holmes, S. 2013. phyloseq: An R package for reproducible interactive analysis and graphic of microbiome census data. *PLoS One* 8:e61217.
- McSpadden Gardener, B. B. M. 2004. Ecology of *Bacillus* and *Paenibacillus* spp. in agricultural systems. *Phytopathology* 94:1252-1258.
- Mercer, C. F., Greenwood, D. R., and Grant, J. L. 1992. Effect of plant and microbial chitinases on the eggs of juveniles of *Meloidogyne hapla* chitwood (Nematoda: Tylenchida). *Nematologica* 38:227-236.
- Nilsson, R. H., Larsson, K., Taylor, A. F. S., Bengtsson-Palme, J., Jeppesen, T. S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F. O., Tedersoo, L., Saar, I., Kõljalg, U., and Abarenkov, K. 2019. The UNITE database for molecular identification of fungi: Handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res.* 47:D259-D264.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlenn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H., Szoecs, E., and Wagner, H. 2019. vegan: Community ecology package. R package version 2.5-6.
- Paradis, E., and Schliep, K. 2018. Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35:526-528.
- Preston, G. M. 2004. Plant perceptions of plant growth-promoting *Pseudomonas*. *Philos. Trans. R. Soc. B* 359:907-918.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., and Glöckner, F. O. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* 41:D590-D596.
- R Core Team. 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Reed, S. E., Greifenhagen, S., Yu, Q., Hoke, A., Burke, D. J., Carta, L. K., Handoo, Z. A., Kantor, M. R., and Koch, J. 2020. Foliar nematode, *Litylenchus crenatae* ssp. *mccannii*, population dynamics in leaves and buds of beech leaf disease-affected trees in Canada and the US. *For. Pathol.* 50:e12599.

- Sapkota, R., and Nicolaisen, M. 2015. High-throughput sequencing of nematode communities from total soil DNA extractions. *BMC Ecol.* 15:3.
- Schlegel, M., Queloz, V., and Sieber, T. N. 2018. The endophytic mycobiome of European ash and sycamore maple leaves—Geographic patterns, host specificity and influence of ash dieback. *Front. Microbiol.* 9:2345.
- Severns, P. M., and Sykes, E. M. 2020. Indicator species analysis: A useful tool for plant disease studies. *Phytopathology* 110:1860-1862.
- Sivakumar, N., Sathishkumar, R., Selvakumar, G., Shyamkumar, R., and Arjunekumar, K. 2020. Phyllospheric microbiomes: Diversity, ecological significance, and biotechnological applications. *Plant Microbiome Sustain. Agric.* 25:113-172.
- Skrzeczowski, L. J., Howell, W. E., Eastwell, K. C., and Cavilleer, T. D. 2001. Bacterial sequences interfering in detection of phytoplasma by PCR using primers derived from the ribosomal RNA operon. *Acta Hort.* 550:417-424.
- Slatko, B. E., Luck, A. N., Dobson, S. L., and Foster, J. M. 2014. *Wolbachia* endosymbionts and human disease control. *Mol. Biochem. Parasitol.* 195:88-95.
- Smart, C. D., Schneider, B., Blomquist, C. L., Guerra, L. J., Harrison, N. A., Ahrens, U., Lorenz, K. H., Seemüller, E., and Kirkpatrick, B. C. 1996. Phytoplasma-specific PCR primers based on sequences of the 16S-23S rRNA spacer region. *Appl. Environ. Microbiol.* 62:2988-2993.
- Taylor, J. E., and Crous, P. W. 2001. Morphological variation and cultural characteristics of *Coniothyrium leucospermi* associated with leaf spot of Proteaceae. *Mycoscience* 42:265-271.
- Usyk, M., Zolnik, C. P., Patel, H., and Levi, M. H. 2017. Novel ITS1 fungal primers for characterization of the mycobiome. *MSphere* 2:e00488-17.
- Wasala, S. K., Brown, A. M. V., Kang, J., Howe, D. K., Peetz, A. B., Zasada, I. A., and Denver, D. R. 2019. Variable abundance and distribution of *Wolbachia* and *Cardinium* endosymbionts in plant-parasitic nematode field populations. *Front. Microbiol.* 10:964.
- White, T. J., Bruns, T., Lee, S. B., and Taylor, J. W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315-322 in: *PCR Protocols: A Guide to Methods and Applications*. M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, eds. Academic Press, San Diego, CA, U.S.A.
- Zhao, B. G., and Lin, F. 2005. Mutualistic symbiosis between *Bursaphelenchus xylophilus* and bacteria of the genus of *Pseudomonas*. *For. Pathol.* 35:339-345.
- Zhao, Y., and Qi, M. 2011. Comparative genomics of *Erwinia amylovora* and related *Erwinia* species—What do we learn? *Genes (Basel)* 2:627-639.