

Oak gall wasps lack conserved microbial symbionts

Authors: Shayan Naeem^{1,2*}, Diana Rojas Guerrero^{1,2}, Piotr Łukasik^{1*}

¹ Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, Kraków, Poland

² Doctoral School, PhD Programme in Biology, Faculty of Biology, Jagiellonian University, Kraków, Poland

*Correspondence: Shayan Naeem shayan.naeem@doctoral.uj.edu.pl; Piotr Łukasik p.lukasik@uj.edu.pl

Abstract

Herbivorous insects frequently rely on microbial partners to exploit their host plants, yet for gall-inducing insect taxa, there were few attempts to comprehensively resolve their microbial associations. To address this gap, we characterised microbial abundance, composition and specificity in three Polish populations of cherry oak gall wasps (*Cynips quercusfolii*), a highly specialized phytophagous lineage. Using quantitative sequencing of bacterial and fungal marker region amplicons, we tested whether these insects harbour conserved microbiota likely associated with gall induction or host specialisation.

Cynips quercusfolii and their associated parasitoid *Torymus auratus* harboured notably low bacterial loads, with a median of ca. 2.28×10^4 16S rRNA gene copies per insect. Their bacterial communities, primarily comprising genera with broad environmental distributions, were highly inconsistent among individuals and populations, and no bacterial lineage was consistently abundant or broadly conserved across the sampled gall wasps. In turn, fungal communities showed a stronger population-specific pattern.

These results indicate oak gall wasps do not maintain abundant, conserved microbial symbionts despite their specialised herbivorous lifestyle. Instead, their bacterial communities appear sparse, variable, and dominated by environmental or transient taxa. This contrasts with symbiont-

dependent lifestyles in many herbivores and suggests gall induction is unlikely to rely on stable bacterial mutualisms.

Keywords:

Microbial symbionts, Microbiome quantification, absolute abundance, 16S rRNA amplicon sequencing, metabarcoding, Cynipidae, Host–microbe interactions

1. Introduction

Plant–insect interactions are among the most complex ecological relationships in terrestrial ecosystems, as they are often involved in highly specialised physiological and developmental responses (Serdo and Degaga 2023). In many of these interactions, microbial symbionts contribute substantially to plant adaptation, contributing to their hosts’ ability to break down recalcitrant compounds, detoxify secondary metabolites, and produce limiting nutrients (Hansen and Moran 2014). In many herbivores, microbes may also influence insect–plant interactions more directly by producing bioactive compounds, including phytohormones, that alter plant developmental pathways (Gupta et al. 2023). They can also affect insects’ defence, reproduction, and other traits broadly relevant to ecological specialisation and adaptation (Feldhaar 2011; Łukasik and Kolasa 2024). However, current knowledge of microbial roles in herbivorous insects remains limited to a set of taxa and functional groups, and the extent to which these patterns generalise across herbivorous lineages remains inadequately explored (Wang et al. 2025).

Gall-inducing insects represent one such underexplored lineage, despite being among the most striking examples of plant manipulation in nature. Gall structures arise when insects redirect host plant development to form specialized organs that provide nutrition and protection for the developing larvae (Ward et al. 2022). Gall morphology is often highly specific to both the inducing insect and the host plant, reflecting tightly coordinated developmental interactions. Although the mechanisms underlying gall induction and formation remain incompletely understood, microbial symbionts have been proposed as potential contributors (Desnitskiy et al. 2023) because certain bacteria produce phytohormones such as cytokinins and auxins that influence plant tissue growth and differentiation (Hammer et al. 2021). Indeed, microbial involvement in plant manipulation has been demonstrated in some insect systems. For example, in the leaf-mining moth *Phyllonorycter*

blancardella, larvae induce nutrient-rich “green islands” in leaves by elevating cytokinin levels, a process dependent on the bacterial symbiont *Wolbachia* (Giron et al. 2007; Kaiser et al. 2010; Body et al. 2013). Whether comparable microbial contributions can occur in insect gall systems remains unclear, but no consistent microbial associations have been identified across independently evolved gall-inducing insect clades (Hammer 2021).

Among gall-inducing cynipid wasps (Hymenoptera: Cynipidae), the oak cherry gall wasp, *Cynips quercusfolii*, is notable for its large, conspicuous galls on the leaves of oaks (*Quercus* spp.) (Fig. 1a,b) and its complex life cycle involving alternating sexual and asexual generations that induce distinct gall forms (Rokas et al. 2002). *C. quercusfolii* was among the model systems for the inconclusive study of the mechanisms of gall induction (Pawłowski et al. 2017), with microbial symbionts specifically postulated as one of the possible mechanisms. But while some *C. quercusfolii* specimens were tested for the presence of *Wolbachia*, with no evidence of infections in the specimens examined (Zhao et al. 2021), information on the broader microbial composition and stability in this species or the broader clade is very limited.

This study aimed to investigate the microbial community structure associated with *Cynips quercusfolii* gall systems to understand the potential role of microbiota in gall development. For three populations from South Poland, we first evaluated the genetic diversity of *C. quercusfolii* and its parasitoid, *Torymus auratus*, and then characterised patterns of bacterial and fungal abundance, diversity, and distribution within gall wasps. Finally, we compared the microbial composition of insect and gall samples to identify shared and distinct bacterial taxa, to understand possible insect–microbe–plant interactions within this specialised gall system.

2. Methods

2.1. Insect sampling and selection

Oak galls induced by *Cynips quercusfolii* were collected in October and November 2025 from three locations in southern Poland - Krzczonów (Population 1; 49.735° N, 19.966° E), Klucze (Population 2; 50.335° N, 19.565° E), and Kraków–Pychowice (Population 3; 50.029° N, 19.891° E) (Fig. 1c; Table S1). Samples were obtained by hand-picking mature oak galls from fallen leaves on the forest floor (Fig. 1b). The galls were transported to the laboratory in ventilated mesh bags and, the following day, carefully dissected to separate insects from surrounding gall tissues. In

most cases, galls contained a single adult wasp morphologically identified as *C. quercusfolii*, but some galls instead contained small white larvae, subsequently identified based on COI barcodes as parasitoid *Torymus auratus*. Individual insects and their corresponding gall tissue samples were preserved separately in 96% ethanol for downstream molecular analyses. When two visually distinct gall tissue regions were seen, both were separated. Some individual *C. quercusfolii* wasps that emerged from galls in a mesh bag and thus were not linked to a particular gall were also processed. In the experiments, we also included three reference insect species to obtain a general context for bacterial abundance and community stability comparisons: a braconid parasitoid *Dacnusa sibirica* (commercial culture), fruit fly *Drosophila melanogaster* (laboratory culture), and the honeybee *Apis mellifera* (laboratory culture) (Table S2).

2.2. DNA extraction, library preparation and sequencing

Samples were processed following a modified protocol described previously (Buczek et al. 2024) and subsequently applied in related insect microbiome studies (Kolasa et al. 2023; Nowak et al. 2025). Insects and their corresponding gall tissue samples were processed individually, as were three reference species.

Individual samples were placed in 2 ml tubes containing lysis buffer and homogenised using ceramic beads, followed by incubation at 55 °C for two hours. Then, a 40 µl aliquot of the homogenate was combined with 2 µl of the solution containing 20,000 copies of plasmids carrying artificial 16S rRNA amplification targets. DNA was subsequently purified using SPRI beads and then used for amplicon library preparation following a two-step PCR approach. During the first PCR step, multiple marker regions were simultaneously amplified, including the bacterial 16S rRNA V4 region, the fungal ITS1 and ITS2 regions, and the insect mitochondrial cytochrome c oxidase I (COI) region. Amplification was performed using primer combinations 515F–806R for 16S rRNA, ITS1F–ITS2 for the fungal ITS1 region, ITS7-ITS4 for fungal ITS2, and BF3–BR2 for insect COI; all template-specific primer sequences were preceded by Illumina adapter tails (Table S3). PCR products were cleaned using SPRI beads and used as templates for the second PCR, during which dual-indexed Illumina adapters were added to each sample. Positive and negative controls were included during different stages of library preparation. The resulting amplicon libraries were verified through agarose gel electrophoresis, pooled based on PCR band

brightness, and sequenced in a highly multiplexed Illumina NextSeq 2000 P2 600-cycle flow cell at Małopolska Centre of Biotechnology of Jagiellonian University.

Initial sequencing generated high-quality 16S rRNA, ITS1, and ITS2 datasets; however, COI barcode sequences were not recovered for several samples. COI amplification for these unsuccessful samples was repeated using the BF3 and fwH2 primer pair (Table S3), following protocols as outlined above, which resulted in barcode recovery for most samples.

2.3. Bioinformatic analysis

Sequence data were processed using a modified bioinformatic workflow described previously (Andriienko et al. 2024; Nowak et al. 2025). Raw reads from each sequencing library were initially separated based on primer sequences into bins corresponding to different target regions. Sequences were processed using a custom pipeline integrating Python scripts with established bioinformatic tools (Nowak et al. 2025). Forward and reverse reads were quality-filtered and merged into contigs using PEAR v0.9.11 (Zhang et al. 2014). The resulting contigs were subsequently dereplicated using VSEARCH v2.31.0 (Rognes et al. 2016) and denoised with UNOISE3 implemented in USEARCH v12.0 (Edgar 2016). Denoising was performed separately for each sequencing library to preserve rare sequence variants that might otherwise be lost during global denoising. Chimeric sequences were identified and removed using USEARCH. Taxonomic classification was performed using the SINTAX algorithm against customised reference databases, including SILVA v138.2 for bacterial 16S rRNA sequences, UNITE 10.0 for fungal ITS datasets, and MIDORI v. 169 for COI. Sequences were clustered at 97% similarity using the UPARSE-OTU algorithm implemented in USEARCH. Data summary tables were generated at both ASV and 97% OTU levels to characterise both fine-scale sequence diversity and broader microbial community composition patterns.

For decontaminating ASV tables, chimeric sequences and non-target reads (chloroplast, mitochondrial and other eukaryotic ASVs) were removed from the 16S rRNA datasets, and spike-in reads were separated for later quantification. We then filtered contamination following a workflow outlined previously, based on the assumption that true experimental sample-derived genotypes should be much better represented in these samples than in negative controls (Łukasik et al. 2017; Surmacz et al. 2024). Relative abundances were recalculated for each library using only the remaining bacterial reads. For each ASV, we then compared its maximum relative

abundance in any experimental library with its maximum relative abundance in any of the negative controls (extraction, PCR, and indexing blanks). ASVs for which this maximum abundance ratio was below 10 were classified as contaminants and excluded from further analyses. ASVs above that threshold were retained as non-contaminants. For further analysis, negative control libraries dominated by oak-derived reads, or likely affected by cross-contamination, were excluded. Non-bacterial ASVs (oak chloroplast and mitochondria, Eukaryota) and ASVs with no reads left in either blanks or samples were also removed and relative abundances for each library were recalculated. Libraries that contained no spike-in reads, or that contained no 16S rRNA reads after contaminant filtering, were removed (Table S7). The original ASV table, before contamination filtering, with reads categorised, is provided as Table S5, and the decontaminated dataset as Table S6. COI data were decontaminated with the same negative-control rule (no ASVs flagged), but still some non-arthropod ASVs and ASVs with fewer than 20 reads were removed (Table S4). ITS1 and ITS2 zOTUs were filtered by removing chimeric sequences, and contaminants identified with the same sample-to-blank ratio rule as for 16S. The original ASV tables for ITS1 and ITS2, before contamination filtering, are provided as Supplementary Tables S11 and S14, and the decontaminated datasets are provided as Supplementary Tables S12 and S15, respectively. Absolute bacterial abundance was then estimated for 16S rRNA datasets using synthetic spike-in standards that enabled converting relative sequencing read counts into absolute abundance, following Tourlousse et al. (2017). Briefly, the ratio of decontaminated bacterial reads to spike-in reads were multiplied by the known number of spike-in copies added before homogenate purification (20,000 copies per one-fifth of the homogenate). Samples lacking spike-in reads were excluded from downstream abundance analyses. Absolute abundance values for the cleaned dataset are provided in Table S7. Reference datasets from *Dacnusa*, *Drosophila*, and *Apis* (Tables S8-10) provide comparative context for bacterial abundance and community composition, as these insects are known to harbour well-established bacterial symbionts consistently detected across individuals. Because no spike-in standards were included for fungal ITS datasets, fungal community analyses were based on relative abundance values only.

Finally, for both bacterial and fungal datasets, OTU abundance tables were generated from ASV abundance matrices. OTUs were subsequently ranked based on average relative abundance across libraries, and the most abundant taxa (top 19–20 OTUs depending on the dataset) were visualised in downstream plots. Bacterial community composition was compared among populations using

Bray–Curtis dissimilarity and permutational multivariate analysis of variance (PERMANOVA) with 999 permutations. Homogeneity of multivariate dispersion was assessed using PERMDISP to evaluate within-population variation. Differentially abundant ASVs were identified using ANCOM-BC2 (Lin and Peddada 2024), with Benjamini–Hochberg correction for multiple testing and a significance threshold of $q < 0.05$.

2.4. Data visualisation

COI sequence clustering and haplotype analyses were performed using MEGA Cluster Explorer (Kumar et al. 2018) and PopART (Leigh et al. 2015). All other visualisations were created in R version 4.5.2. Heatmaps were generated using the ComplexHeatmap package (Gu et al. 2016) with additional visualization support from circlize (Gu et al. 2014) and grid (Murrell 2002). Microbial community composition was further explored using Non-metric Multidimensional Scaling based on Bray–Curtis dissimilarity metrics. NMDS analyses were performed using the phyloseq package (McMurdie and Holmes 2013), while ecological distance calculations were conducted using vegan (Dixon 2003). Additional graphical visualizations, including abundance boxplots, were generated using ggplot2 (Wickham 2016).

3. Results

3.1. The diversity of the sampled gall wasps

We obtained reliable COI amplicon data for 57 wasps individually extracted from oak galls collected from three locations in Southern Poland (Fig. 1c), including 51 adults that were morphologically identified as *Cynips quercusfolii*, and six parasitoid larvae. For the adult wasps, we obtained between 206 and 4653 COI reads per library, with a single dominant ASV in each individual, considered as its barcode. These 51 barcodes represented 8 distinct haplotypes, separated by up to 3 nucleotide positions and with the consensus haplotype most abundant. All were confidently taxonomically assigned as *C. quercusfolii* (Fig. 1d; Table S4). The COI data for the six larvae yielded between 97 and 4635 reads per library. Their barcodes represented five distinct haplotypes, separated by up to 5 nucleotide positions (Fig. 1e; Table S4), but all were confidently identified as *Torymus auratus* (Hymenoptera: Torymidae) – a known parasitoid of *C. quercusfolii*.

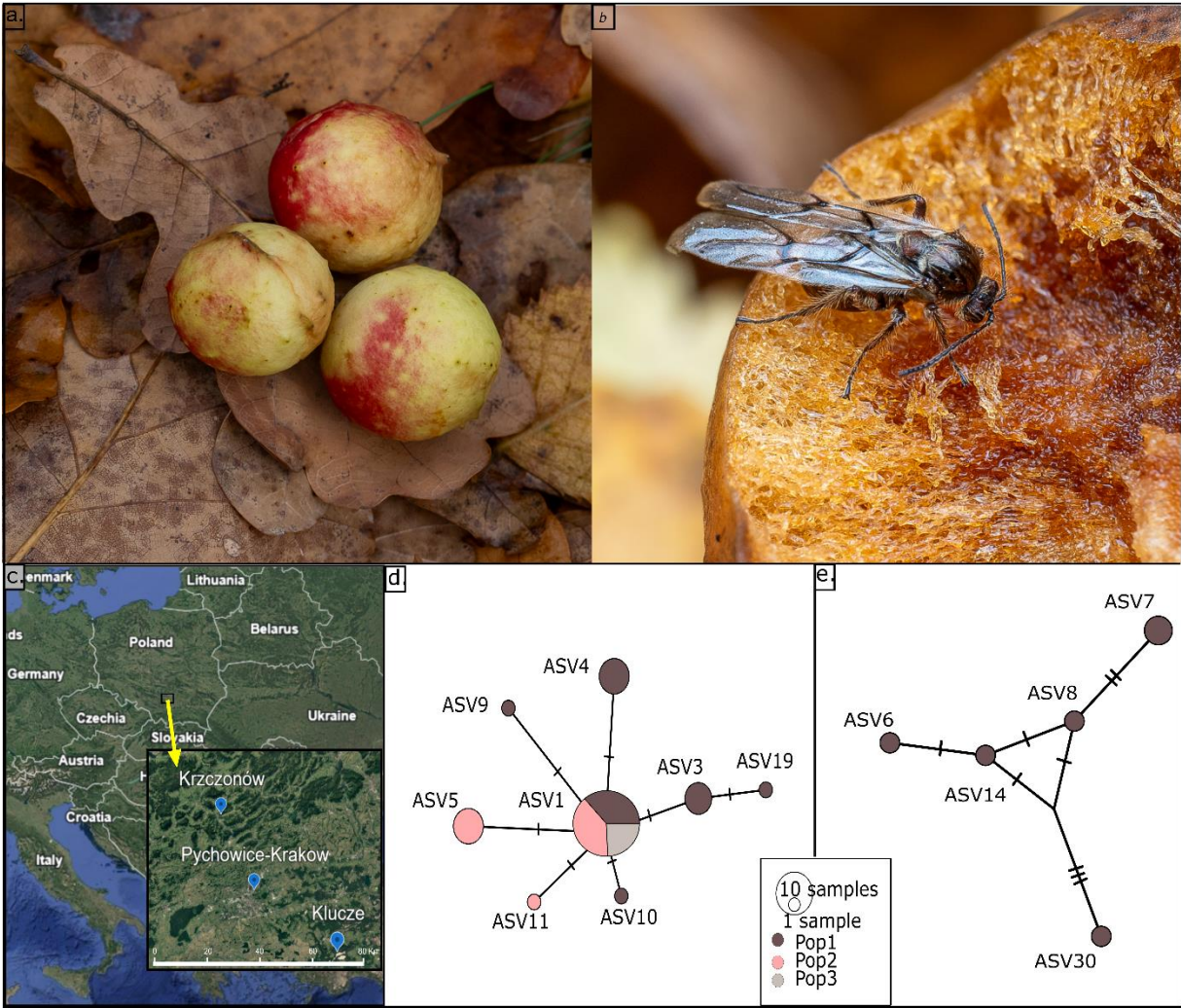


Fig. 1. (a) Oak galls induced by *Cynips quercusfolii* at the time of collection, in October 2025. (b) Adult *C. quercusfolii* on gall tissue. Photos: P. Ł. (c) Three gall wasp sampling locations in southern Poland: Klucze (Pop. 2), Kraków: Pychowice (Pop. 3), and Krzczonów (Pop. 1). (d) Haplotype network for 51 *Cynips quercusfolii* specimens, showing multiple COI genetic variants (ASVs) across populations. Circle size is proportional to haplotype frequency; colours indicate populations, and small bars over the lines represent mutational steps. (e) Haplotype network for six individuals of the parasitoid *Torymus auratus*, showing multiple genetic variants across samples. Panels (d) and (e) use the same colour scheme.

3.2. Composition of 16S rRNA reads and contaminant filtering

The 16S rRNA amplicon data for wasps and their galls were dominated by non-bacterial reads, including chloroplasts, mitochondria, and quantification spike-ins (Fig. 2a). Specifically, of 7,369,006 total reads, oak chloroplast reads made up 52.0% and oak mitochondrial reads 16.4%, so that about two-thirds of all reads were plant-derived. Spike-in standards accounted for 15.6%, and all other reads, comprising bacteria, archaea and eukaryotes, for 16.0% (Fig. 2a; Table S5). The composition differed strongly between sample types, with plant-derived sequences particularly common in gall tissue libraries, while gall wasp and parasitoid libraries contained a much smaller and variable plant-derived fraction and a larger share of spike-in reads. Reference species libraries consisted almost entirely of bacterial reads, with little oak or spike-in signal.

Based on comparisons among experimental samples and controls, wasp and gall samples contained a substantial share of apparent reagent-derived contaminants. Negative control-based contamination filtering removed 25.5% of bacterial reads, on average, from libraries representing these samples (range 0-100%) – much more than from “reference” species (Fig. 2b). The most abundant contaminant, ASV13 (*Polaromonas*), accounted for 22.6% of the summed relative abundance of all contaminant ASVs, whereas the most abundant non-contaminant, ASV19 (*Lichenibacterium*), accounted for 8.6% of the summed relative abundance of all non-contaminant ASVs. Contamination was minor in both the ITS1 and ITS2 data; raw and cleaned datasets differed little, as filtering removed mainly chimaeras and very few contaminants (Table S11-16).

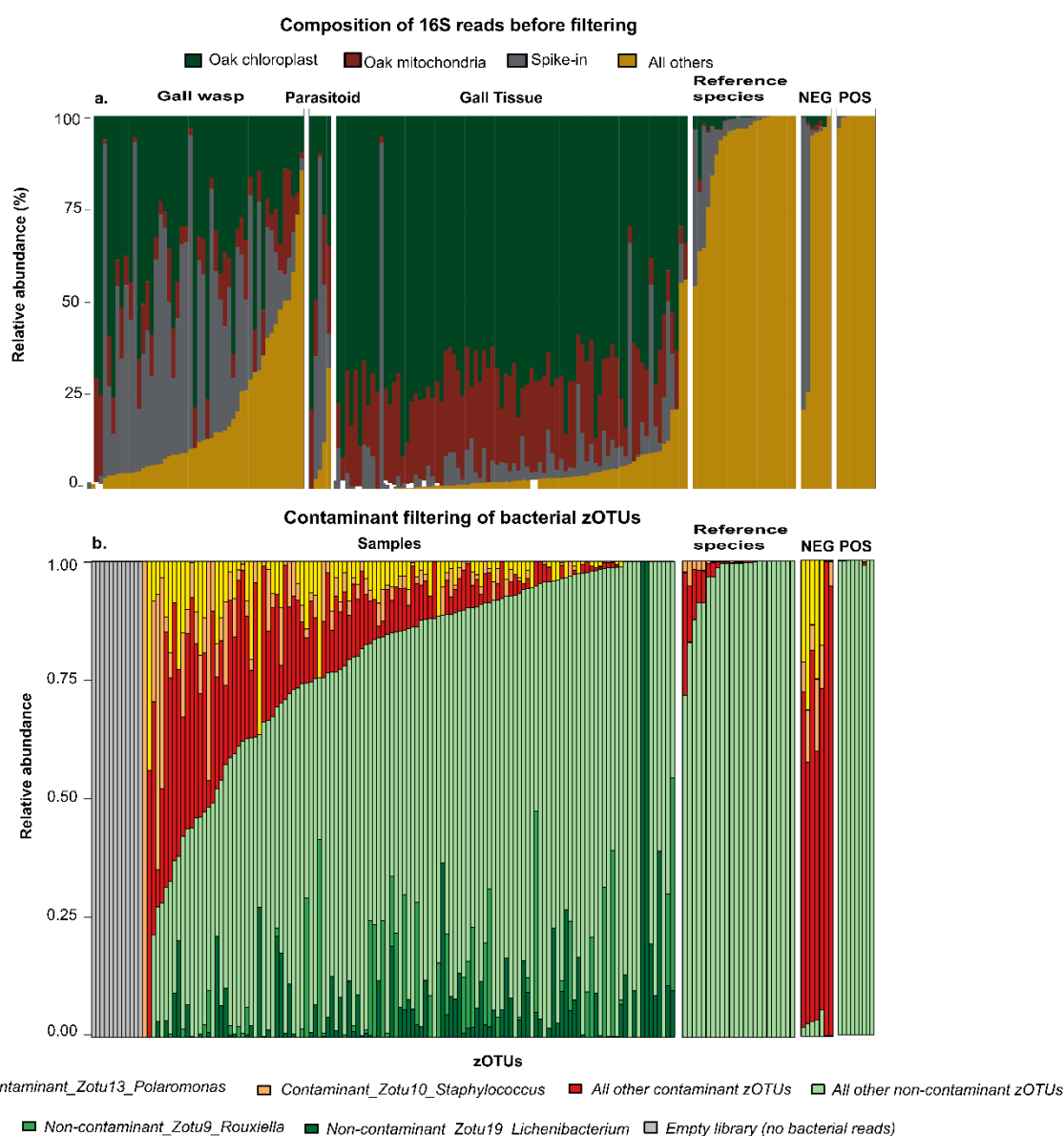


Fig. 2. (a) Relative abundance of oak chloroplast reads, oak mitochondrial reads, spike-in reads, and all other taxonomic categories combined in each library before filtering. Libraries, represented by bars, are grouped by sample type (gall wasp, parasitoid, gall tissue, reference species, negative controls [NEG] and positive controls [POS]) and sorted within each group by increasing proportion of “other” reads. (b) Relative abundance of contaminant and non-contaminant bacterial ASVs in each library after removal of non-target reads. Libraries are grouped into Samples (wasps and galls), Reference species, negative controls (NEG) and positive controls (POS), and sorted by

decreasing proportion of apparent contaminants. The two most abundant contaminant ASVs (ASV13, *Polaromonas*; ASV10, *Staphylococcus*) and the two most abundant non-contaminant ASVs (ASV19, *Lichenibacterium*; ASV9, *Rouxiella*) are shown separately, and all remaining contaminant and non-contaminant ASVs are grouped. Libraries without bacterial reads after filtering are shown in grey.

3.3. Low absolute bacterial abundance in gall wasps

After contamination filtering, we obtained 16S rRNA data that we deemed reliable for 48 of the 57 processed wasps. (Table S6). The number of reads corresponding to non-contaminant 16S rRNA reads per library ranged from 11 to 20,321 reads (median: 2,316 reads), whereas the number of corresponding quantification spike-ins, of which we added ca. 20,000 copies per one-fifth of the whole-insect homogenate, was typically several-fold greater, ranging from 11 to 122,980 reads (median: 8,581 reads) (Table S6). Based on the spike-in to non-contaminant read ratio, estimated bacterial abundance ranged from approximately 3.55×10^3 to 1.56×10^8 16S rRNA gene copies per sample in *Cynips quercusfolii* (median: 2.44×10^4 copies), and from approximately 3.60×10^3 to 1.65×10^5 copies per sample in *Torymus auratus* (median: 1.30×10^4 copies). (Fig. 3a; Table S7).

These estimated 16S rRNA copy numbers were, on average, at least an order of magnitude less than in the three other insect species used as references in this study. In the braconid wasp *Dacnusa sibirica*, the estimated 16S rRNA gene copy numbers ranged between 1.04×10^5 and 4.24×10^6 per insect (median: 4.15×10^5), in the fly *Drosophila melanogaster* between 2.77×10^6 and 8.79×10^6 copies (median: 3.25×10^6), and in the honeybee *Apis mellifera*, between 1.87×10^7 and 3.12×10^8 copies (median: 1.06×10^8) (Fig. 3a; Table S10).

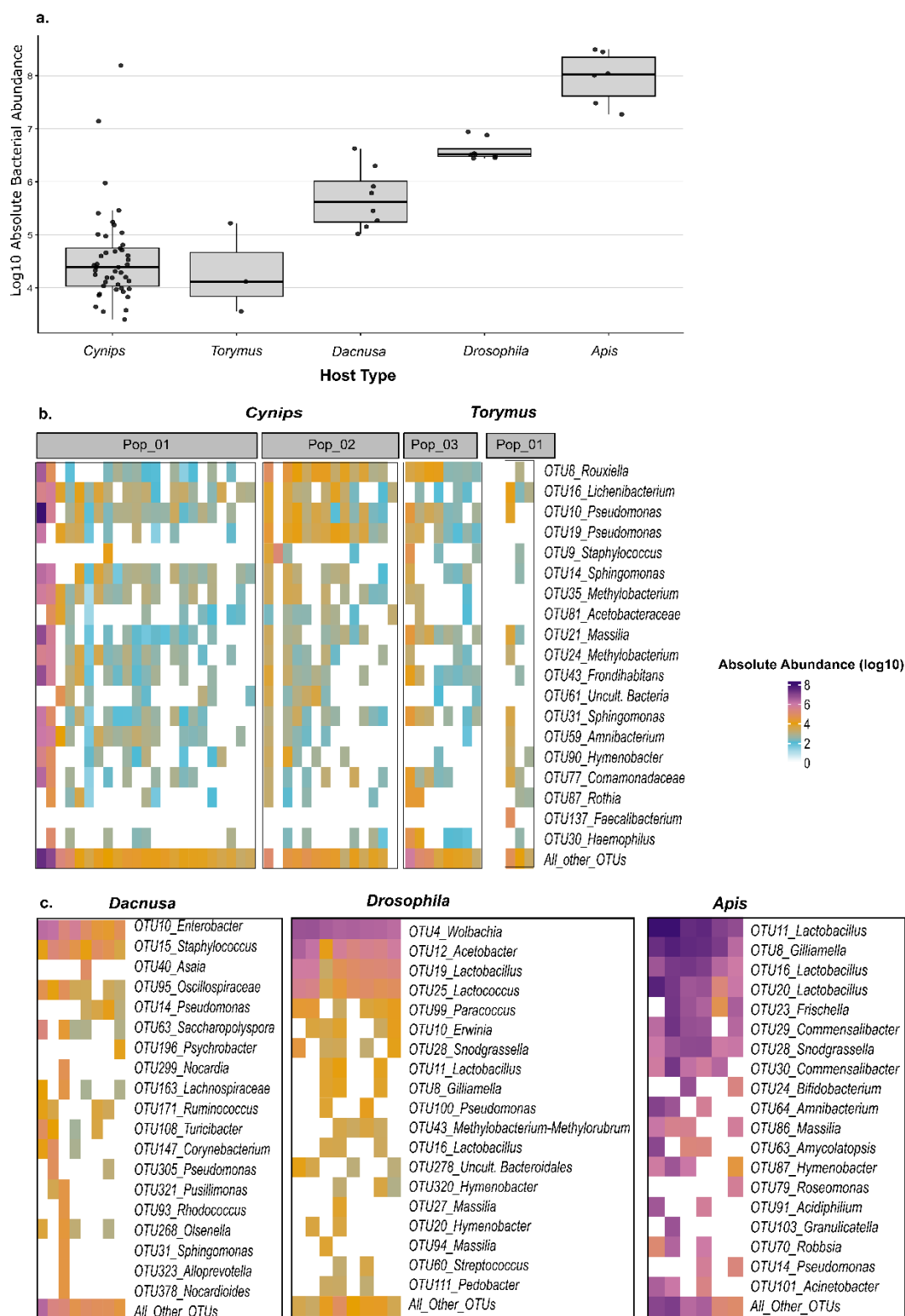
3.4. Inconsistent bacterial clade distribution across wasp populations

Across the 48 processed gall wasp samples, we identified 2,677 16S rRNA non-contaminant amplicon sequence variants (ASVs). Their clustering into 1,239 97% Operational Taxonomic Units (OTUs) facilitated comparisons among samples. These OTUs represented a diverse set of bacterial genera, although their relative abundances were unevenly distributed across the community. The most common genera, based on average relative abundance across all gall wasps, were *Pseudomonas* (9%), followed by *Rouxiella* (7%), *Lichenibacterium* (6%), *Sphingomonas*

(4%), *Staphylococcus* (3%), *Methylobacterium* (3%), *Massilia* (1%), *Fronidhabitans* (1%), *Amnibacterium* (1%), *Hymenobacter* (1%), *Rothia* (1%), *Faecalibacterium* (1%), and *Haemophilus* (1%). Overall, the OTUs representing the genera listed above made up approximately 45% of the microbial communities on average (Table S7; Fig. 3b). Several of the dominant OTUs were also widely distributed across the gall wasp samples. *Pseudomonas* OTU10 was detected in 39 of the 48 samples (81.3%), making it one of the most widespread OTUs, followed by *Lichenibacterium* OTU16, which occurred in 36 samples (75.0%), and *Rouxiella* OTU8, detected in 35 samples (72.9%). No single OTU was consistently present across all gall wasp individuals, or all individuals in a population. Further, taxon-specific bacterial abundances were relatively low and showed no consistent patterns across populations. OTUs that reached relatively high abundances in individual samples were not consistently abundant across the *Cynips* populations or in *Torymus auratus* individuals.

To formally assess differences among *Cynips* populations, we compared their bacterial community composition using PERMANOVA. Community composition differed significantly among the three populations ($F_{2,42} = 2.15$, $R^2 = 0.093$, $p = 0.001$; Table S17a), with all three pairwise comparisons remaining significant after BH correction (Table S17b). PERMDISP showed no significant differences in within-population dispersion ($F_{2,42} = 1.21$, $p = 0.310$; Table S17c), indicating that the PERMANOVA result was not explained by differences in dispersion. At the individual OTU level, global ANCOM-BC2 identified three OTUs (OTU8, OTU19, and OTU59) with significant differences in abundance among populations (Table S18). However, in pairwise ANCOM-BC2 comparisons, only OTU_59 showed a significant difference between populations 1 and 3 (LFC = -1.73, $q = 0.029$; Table S19).

The variation within host species contrasted with that in the three cultured reference species that showed much more consistent associations (Fig. 3c). For example, the microbiota of the cultured *Drosophila* was consistently dominated by *Wolbachia*, *Acetobacter*, and *Lactobacillus*. Similarly, honeybees hosted their characteristic core taxa (Fig. 3c; Table S9-10).



298 **Fig. 3.** (a) Estimated absolute numbers of 16S rRNA copies ($\log_{10}$ scale) in individuals of *Cynips*
 299 *quercusfolii*, its parasitoid *Torymus auratus* and in three reference species (*Dacnusa sibirica*,
 300 *Drosophila melanogaster*, and *Apis mellifera*). (b) The distribution and estimated absolute
 301 abundance of the 20 most common bacterial OTUs across gall wasps (*C. quercusfolii*) from three
 302 populations and in the associated parasitoid *Torymus auratus*. Each row represents a bacterial
 303 OTU, and each column represents an individual insect. Colors indicate an estimated absolute
 304 abundance of a specific OTU in the specific insect. (c) The distributions and estimated absolute
 305 abundances of dominant bacterial OTUs in three reference insect species. The color scale in this
 306 panel is the same as in (b).

307 3.5. Overlap in bacterial communities between wasps and gall 308 tissues

309 Similarly to the gall wasps, the bacterial community associated with their respective galls was
 310 highly variable across samples and populations. Most ASVs and OTUs were only present in a
 311 subset of samples, resulting in a patchy community structure (Fig. 4a; Table S6-7). Although many
 312 OTUs present in galls overlap with those seen in insects, this overlap is inconsistent and not
 313 maintained across all populations. As seen for the wasps, gall tissue samples harbor distinct
 314 combinations of OTUs, and no dominant taxa are seen. Overall, the community appears to be
 315 largely composed of rare and sample-specific taxa and provides limited evidence for structured or
 316 conserved associations between insects and their gall tissues (Fig. 4a).

317 Ordination analysis (Fig. 4b) further supports this pattern of variability but also suggests some
 318 structuring at the population level. The spatial proximity between insects and their corresponding
 319 gall samples from the same population share more similar bacterial assemblages compared to those
 320 from other populations. To test whether insects harboured bacterial communities more similar to
 321 their own associated galls than to non-associated galls, Bray–Curtis dissimilarities were compared
 322 between matched and non-matched insect–gall pairs. Matched insect–gall pairs showed
 323 significantly lower Bray–Curtis dissimilarities than non-matched pairs (Wilcoxon signed-rank test,
 324 $V = 177$, $P = 0.0326$; Table S20). Overall, bacterial community composition also differed between
 325 insects and galls (PERMANOVA, $R^2 = 0.0358$, $F = 3.046$, $P = 0.001$), while PERMDISP indicated
 326 no significant difference in within-group dispersion ($F = 2.861$, $P = 0.095$; Table S20). To further

determine which populations accounted for these differences, pairwise ANCOM-BC2 comparisons were performed between populations 1, 2 and 3, while accounting for sample type (Insect vs Gall) (Table S21). Significant population-specific differences were detected across all three pairwise comparisons, with 22 OTUs differing in at least one comparison (Table S22). Several taxa, including *Lichenibacterium*, *Pseudomonas*, *Rouxiella*, *Amnibacterium*, *Hymenobacter*, *Rothia*, *Spirosoma*, and *Ovatusbacter*, showed population-associated differences. Four OTUs in the population 1 and 2 comparison additionally met the robust differential-abundance criterion, whereas no OTUs met this criterion in the population 1 and 3 or 2 and 3 comparisons.

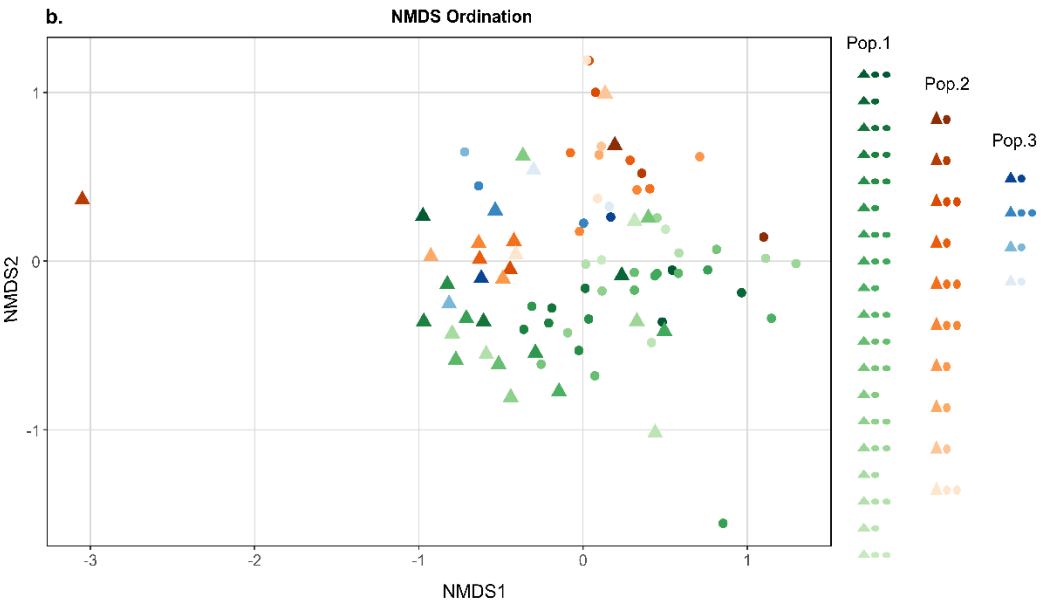
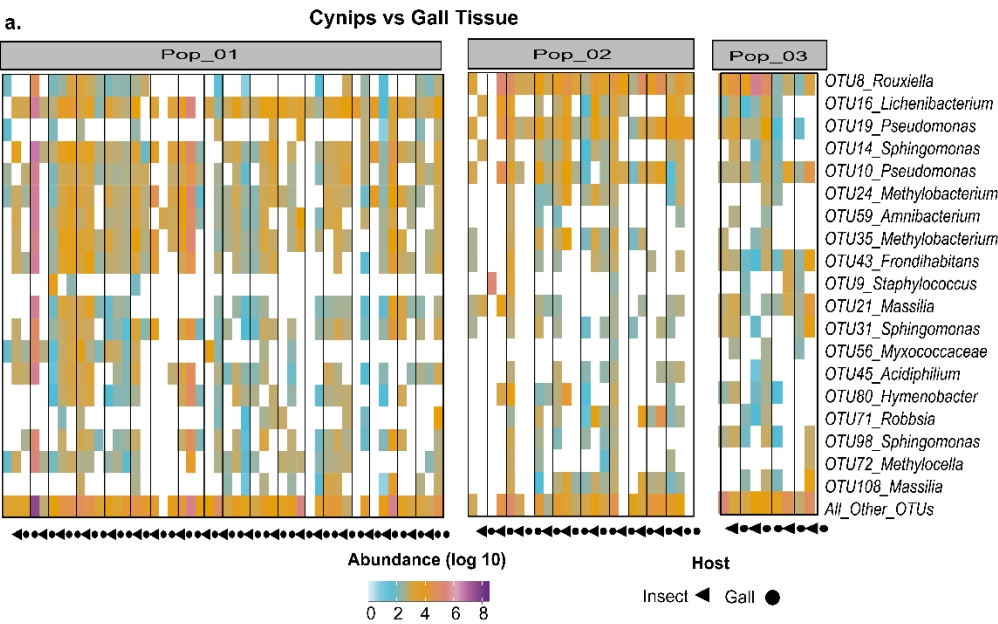


Fig. 4. (a) The distribution of most common bacterial OTUs in *Cynips* wasps from three populations and in their corresponding gall tissues. Each row represents a bacterial OTU, and each column corresponds to an individual insect, or a gall tissue sample, with wasps and their corresponding galls separated by vertical lines. Color intensity indicates the absolute abundance of 16S rRNA copies ($\log_{10}$ scale). (b) Non-metric multidimensional scaling (NMDS) plot of bacterial community structure in insect and gall samples. Points represent individual samples; shapes separate host type, and colors represent sets of insect-galls within the population.

3.6. Site-specific fungal communities in *Cynips* wasps and their gall tissues

Fungal communities across *Cynips* wasps and their associated gall tissues, assessed based on ITS1 and ITS2 region data, again showed variability among individuals and populations (Fig. 5a-b). No quantification spike-ins for ITS regions were used; hence our conclusions are based on relative abundance information only.

The ITS1 clean data ranged from 8 to 77,151 reads per sample (mean = 7,012, median = 3,006) (Table S2 and S12). In total, we identified 1,185 ASVs, which clustered into 710 OTUs, with OTU richness ranging from 6 to 122 OTUs per sample (Table S12-13). The most abundant fungal taxa included *Microcyclospora*, *Ramularia*, *Curvibasidium*, *Magnohelicospora*, *Epicoccum*, *Phaeotremella*, *Suillus*, *Xenosonderhenia*, *Aureobasidium*, *Cladosporium*, and *Vishniacozyma* (Fig. 5a). Relative abundances differed markedly among samples, with some taxa occurring at high abundance in specific individuals while being absent in others. However, we observed consistent differences among populations. For example, OTU2_*Microcyclospora* and OTU4_*Curvibasidium* were both consistently detected and abundant across samples from population 1, where they occurred in both insects and gall tissues, but were largely absent from the remaining two populations. In turn, in Population 2, the dominant fungal clade was OTU3_*Ramularia*, less common in other populations.

ITS2 amplicon data ranged from 42 to 139,983 clean reads per sample (median = 2,885, mean = 6,934) (Table S2 and S15). In total, we identified 1360 ASVs, which clustered into 808 OTUs, with 1–153 OTUs detected per sample (Fig. 5b; Table S15-16). The dominant fungal genera included *Ramularia*, *Colletotrichum*, *Magnohelicospora*, *Curvibasidium*, *Vishniacozyma*, *Nothophoma*, *Aureobasidium*, *Cyanodermella*, *Ciboria*, *Biatora*, *Candida*, and *Peltaster*. As in the

ITS1 dataset, fungal community composition was highly heterogeneous, with no fungal group consistently associated with all insect or gall samples, but again with clear population-level differences paralleling those observed in ITS1 dataset. For example, OTU representing *Ramularia* was highly abundant in Population 2, and *Curvibasidium* in Population 1. On the other hand, *Microcyclospora* was not among the abundant clades captured by ITS2 primers.

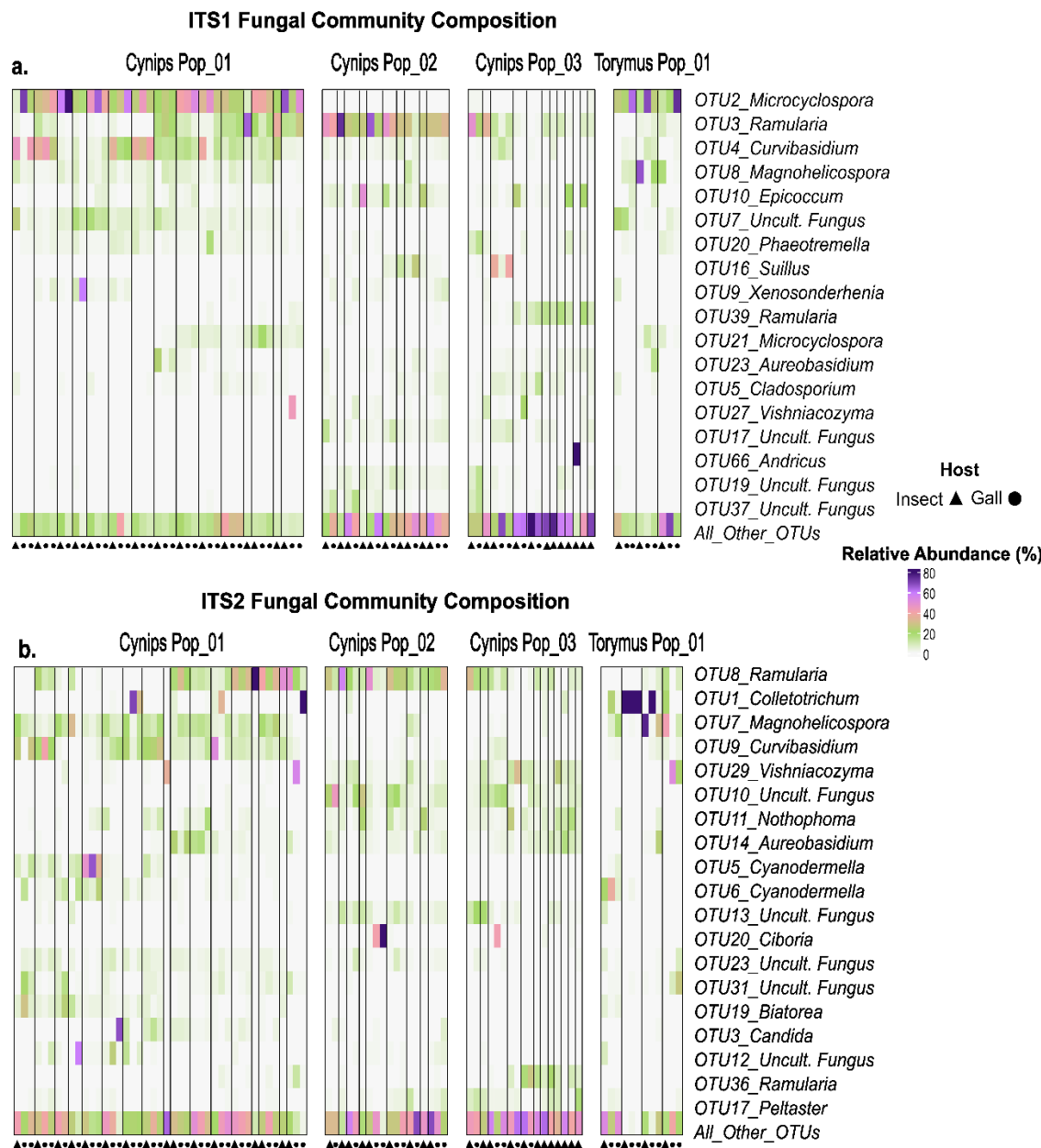


Fig. 5. (a) The relative abundance (%) of fungal OTUs identified through ITS1 sequencing in insect hosts and associated gall tissues across *Cynips* (Pop_1–3) and *Torymus* (Pop_1) populations.

(b) The relative abundance (%) of fungal OTUs identified through ITS2 sequencing in insect hosts and associated gall tissues. Columns represent individual samples, and rows represent fungal OTUs. Color intensity reflects relative abundance.

3.7. Fungal Community Composition Varies Among Populations and Sites

Fungal communities showed clear differences among populations and sampling sites in both the ITS1 and ITS2 datasets. The NMDS plots showed a clear separation of Population 1 from Populations 2 and 3 (Fig. 6a-b). The separation of Population 1 from Populations 2 and 3 suggests that fungal assemblages were not uniformly distributed across the sampled populations, and the greater overlap between populations 2 and 3 indicates greater similarity in their fungal communities. PERMANOVA revealed significant differences in fungal community composition among the three collection sites for both ITS1 ($F_{2,70} = 20.80$, $R^2 = 0.373$, $p = 0.001$) and ITS2 ($F_{2,75} = 9.62$, $R^2 = 0.204$, $p = 0.001$) datasets. Pairwise comparisons confirmed significant differences between all three site pairs (all $p \leq 0.001$). These differences were accompanied by significantly greater within-site dispersion in Pop_03 for ITS1, whereas dispersion did not differ significantly among sites for ITS2 (Table S23).

Overall, these differences among populations suggest that the fungi associating with galls and gall wasps may be primarily shaped by the local environment rather than representing a specialised fungal consortium required for gall formation.

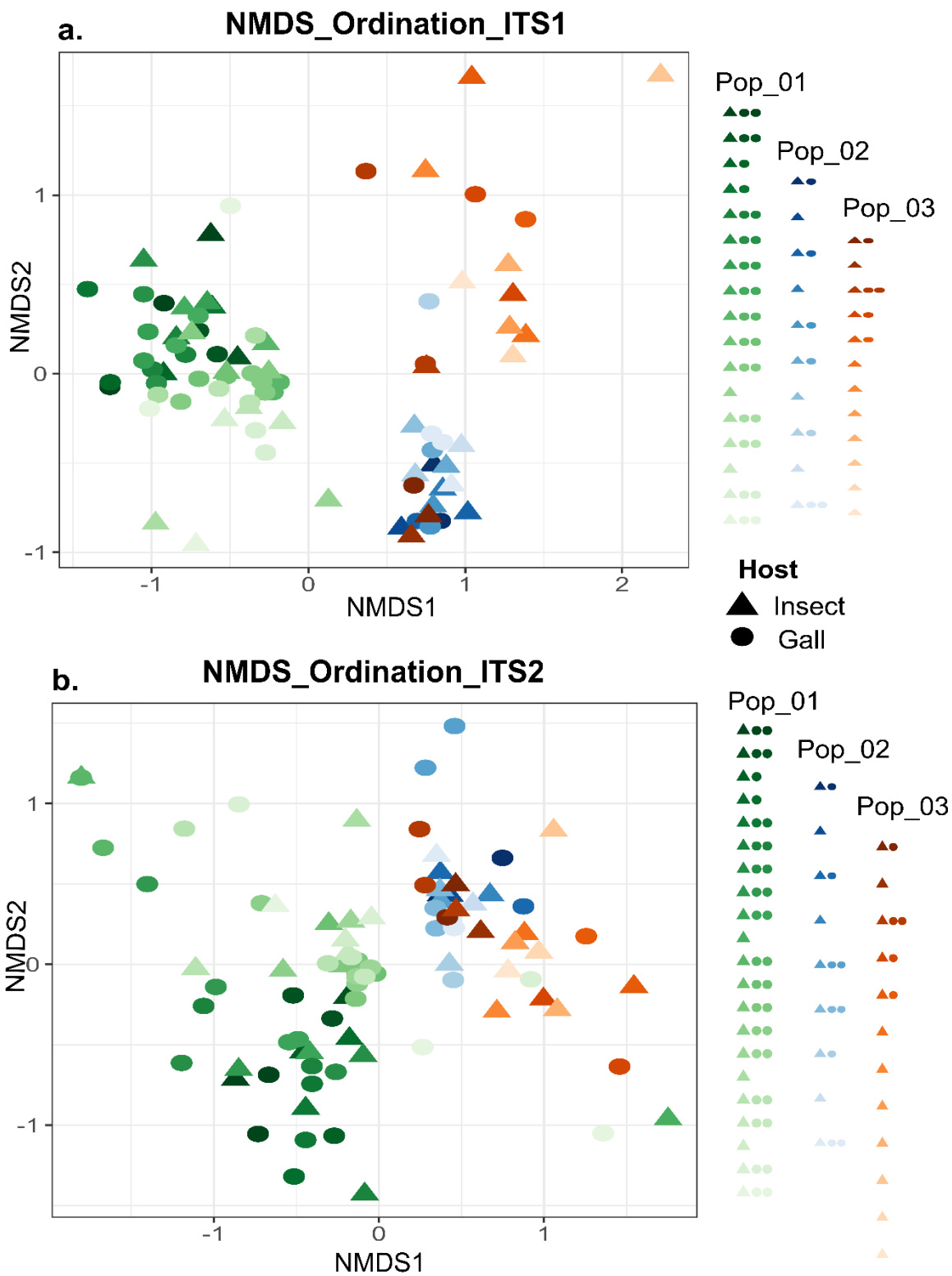


Fig. 6. (a) NMDS ordination based on ITS1 sequences. (b) NMDS ordination based on ITS2 sequences. Symbols represent host types, with triangles indicating insects and circles indicating galls. Colors indicate the three sampled populations (Pop_01, Pop_02 and Pop_03).

4. Discussion

Multi-marker amplicon sequencing combined with microbiota quantification is an effective tool for investigating the microbiota of wild-collected samples with variable but often low bacterial load (Nowak et al. 2025). *Cynips quercusfolii* and their corresponding gall tissues represent such case. Here, by combining COI barcode reconstruction with bacterial 16S rRNA and fungal ITS metabarcoding, we assessed host-associated genetic diversity together with bacterial and fungal community composition and – for bacteria – their abundance. These combined data allowed us to resolve microbial patterns across insect hosts, gall tissues, populations, and some associated parasitoids to understand microbiome variation within oak gall systems better.

Before examining microbial community composition, it is important to consider the genetic diversity of the host insects, as cryptic diversity and host evolutionary history can shape microbiome structure (Ward et al. 2022; Nowak et al. 2025). In our study, COI amplicon sequencing revealed multiple ASVs within both *Cynips quercusfolii* and the associated parasitoid *Torymus auratus*, adding to evidence of fine-scale mitochondrial variation in gall-associated insect systems. Previous studies have already demonstrated pronounced genetic structuring among geographically separated *Cynips* populations in Turkey, Iran, Hungary, and Spain (Mutun and Dinç 2023), or in other oak-associated gall wasps or their parasitoids (Mutun et al. 2019). COI amplicon data help resolve that diversity.

The bacterial communities of *Cynips quercusfolii* and its parasitoid *Torymus auratus* were of low abundance and highly variable across individuals, populations, and associated gall tissues, with no evidence of a stable or universally conserved bacterial symbiont. Low spike-in-based abundance estimates were concordant with a high share of plant-derived reads and contaminant genotypes in these samples. A similar pattern of low bacterial abundance has been reported in many other insects (Hammer et al. 2019). For example, Hammer (2017), using quantitative PCR and 16S rRNA gene sequencing across 124 species of larval Lepidoptera, showed that bacterial densities in caterpillar

guts were orders of magnitude lower than those observed in other insects and vertebrates with known resident microbiomes, while Sanders (2017) demonstrated that many clades of ants lack abundant, specialized microbiota. Further, gall wasps primarily associated with bacterial genera broadly distributed in the environment, including *Pseudomonas*, *Massilia*, and *Sphingomonas*, none of which was consistently detected across all wasp individuals or gall tissue samples. This argues for the transience and environmental source of bacteria detected within *Cynips* wasps. Gall wasps also did not appear to harbor putative obligate or facultative endosymbionts: few *Wolbachia*, *Rickettsia*, or *Spiroplasma* ASVs were all sparsely distributed and of low abundance (Table S6), suggestive of an environmental source or perhaps laboratory contamination. Hammer (2021) argued that if bacterial symbionts were essential for gall induction - an obligate trait in these insects - they would be expected to occur consistently across host individuals. This is not the case here.

The patterns in *C. quercusfolii* contrast with reference insects such as *Apis*, *Drosophila*, and *Dacnusa*, which consistently harboured several bacterial taxa, of much greater abundances than the gall wasps. The difference among these systems supports the interpretation that abundant, stable bacterial microbiota are not a general insect feature (Hammer et al. 2019; Ravenscraft and Coon 2026). However, transient microbes – as those apparently seen in *Cynips* - may persist in insects for extended periods, and importantly, transience does not necessarily imply biological insignificance (Ravenscraft & Coon 2026).

The role of fungi in gall formation has also been proposed (Rohfritsch 2008; Raman et al. 2012). Although we found no single fungal group that dominated across all host populations, several fungal taxa showed population-specific patterns. Geographic variation in gall-associated fungal communities has also been reported in other cynipid systems. In *Dryocosmus kuriphilus*, fungal community composition differed between geographically separated regions and sampling sites (Meyer et al. 2015). These differences may largely reflect environmental, temporal, or even stochastic variation in oak endophyte communities across collections. Similar site-specific patterns in associated fungal communities have also been reported in forest ants (Siedlecki et al. 2026). Such patterns in galls and ants suggest that insect-associated fungal communities may partly reflect environmental acquisition from the surrounding forest rather than stably associating with insect hosts.

The absence of a consistently dominant fungal group across our *C. quercusfolii* populations also raises the possibility that many of these fungi represent secondary colonisers rather than stable gall-associated symbionts. (Batra and Lichtwardt 1963) observed that fungal occurrence within galls was highly variable, nonspecific, and frequently associated with older or damaged gall tissues rather than with early developmental stages. Washburn and Van Bael (2017) found diverse fungal communities inside baldcypress galls, but no conserved fungal symbiont linked to gall initiation. Instead, fungal composition varied with insect emergence status, and fungi were proposed to colonize galls after formation either from surrounding plant tissues or during oviposition.

Nevertheless, fungi can play a more direct functional role in some gall-forming systems. Rohfritsch (2008) described ambrosia galls produced by gall midges in which fungal mycelium develops within the gall and serves as a food source for developing larvae. Thus, although fungi may form an integral component of some gall systems, the population-specific associations observed in our study do not by themselves demonstrate a role in gall initiation.

All our above findings are consistent with evidence that the induction of cherry oak galls is mediated directly by insect-derived molecular mechanisms, rather than bacterial symbionts. Indeed, genomic and transcriptomic studies of gall-inducing Cynipidae have revealed the overexpression of genes encoding plant cell wall-degrading enzymes and venom-like proteins during gall development (Mozhaitseva 2023). Similarly, recent molecular studies on insect gall formation have demonstrated that gall-inducing insects manipulate host plant hormonal, developmental, and metabolic pathways through insect-secreted effectors, including phytohormone-like compounds, salivary proteins, and other regulatory molecules (Takeda et al. 2021). Raman (2021) proposed that gall induction is initiated through mechanical injury caused by the insect, with the release of insect-derived effectors and chitin-associated signals that trigger localized cellular reprogramming in host tissues. That study emphasized that no consistent third-party organisms or virus-like particles were detected in various Cynipid gall systems.

Overall, our study shows that microbial communities associated with *Cynips quercusfolii* gall systems are low in abundance, highly variable, and lack evidence of stable and conserved symbiotic associations across individuals and host populations. Both bacterial and fungal communities showed inconsistent distributions among insects and gall tissues, with no single microbial taxon consistently associated with all gall samples. This combination of low abundance

and inconsistent community composition argues against a consistent microbial role in gall formation. Instead, the observed microbiota appears to be largely influenced by local environmental conditions, host-specific factors, and secondary colonization processes within already developed gall tissues.

5. Acknowledgements

This study was supported by the Polish National Science Centre grant 2021/43/B/NZ8/03376 and by a grant from the Priority Research Area BioS under the Strategic Programme Excellence Initiative at Jagiellonian University.

6. Contributions

Shayan Naeem: Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Piotr Łukasik:** Conceptualization, Validation, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Diana Rojas Guerrero:** Investigation (sample collection), Validation, and Supervision.

7. Data Availability Statement

The supplementary data that support the findings of this study are openly available in figshare at <https://doi.org/10.6084/m9.figshare.34704399>. Raw amplicon sequencing data are available through NCBI BioProject PRJNA1544549.

8. References

- Andriienko V, Buczek M, Meier R, et al (2024) Implementing high-throughput insect barcoding in microbiome studies: impact of non-destructive DNA extraction on microbiome reconstruction. *PeerJ* 12:e18025
- Batra LR, Lichtwardt RW (1963) Association of fungi with some insect galls. *J Kans Entomol Soc* 36:262–278
- Body M, Kaiser W, Dubreuil G, et al (2013) Leaf-miners co-opt microorganisms to enhance their nutritional environment. *J Chem Ecol* 39:969–977
- Buczek M, Prus-Frankowska M, Łukasik P (2024) Quantitative multi-target amplicon sequencing workflow

518 Desnitskiy AG, Chetverikov PE, Ivanova LA, et al (2023) Molecular aspects of gall
519 formation induced by mites and insects. *Life* 13:1347

520 Dixon P (2003) VEGAN, a package of R functions for community ecology. *J Veg Sci*
521 14:927–930

522 Edgar RC (2016) UNOISE2: improved error-correction for Illumina 16S and ITS
523 amplicon sequencing. *BioRxiv* 081257

524 Feldhaar H (2011) Bacterial symbionts as mediators of ecologically important traits of
525 insect hosts. *Ecol Entomol* 36:533–543

526 Giron D, Kaiser W, Imbault N, Casas J (2007) Cytokinin-mediated leaf manipulation by
527 a leafminer caterpillar. *Biol Lett* 3:340–343

528 Gu Z, Eils R, Schlesner M (2016) Complex heatmaps reveal patterns and correlations in
529 multidimensional genomic data. *Bioinformatics* 32:2847–2849

530 Gu Z, Gu L, Eils R, et al (2014) “CircIzize” implements and enhances circular
531 visualization in R

532 Hammer TJ, De Clerck-Floate R, Tooker JF, et al (2021) Are bacterial symbionts
533 associated with gall induction in insects? *Arthropod-Plant Interact* 15:1–12

534 Hammer TJ, Janzen DH, Hallwachs W, et al (2017) Caterpillars lack a resident gut
535 microbiome. *Proc Natl Acad Sci* 114:9641–9646

536 Hammer TJ, Sanders JG, Fierer N (2019) Not all animals need a microbiome. *FEMS*
537 *Microbiol Lett* 366:fnz117

538 Hansen AK, Moran NA (2014) The impact of microbial symbionts on host plant
539 utilization by herbivorous insects. *Mol Ecol* 23:1473–1496

540 Kaiser W, Huguet E, Casas J, et al (2010) Plant green-island phenotype induced by
541 leaf-miners is mediated by bacterial symbionts. *Proc R Soc B Biol Sci* 277:2311–
542 2319

543 Kolasa M, Kajtoch Ł, Michalik A, et al (2023) Till evolution do us part: The diversity of
544 symbiotic associations across populations of *Philaenus spittlebugs*. *Environ*
545 *Microbiol* 25:2431–2446

546 Kumar S, Stecher G, Li M, et al (2018) MEGA X: molecular evolutionary genetics
547 analysis across computing platforms. *Mol Biol Evol* 35:1547–1549

548 Leigh JW, Bryant D, Nakagawa S (2015) POPART: full-feature software for haplotype
549 network construction. *Methods Ecol Evol* 6:

550 Lin H, Peddada SD (2024) Multigroup analysis of compositions of microbiomes with
551 covariate adjustments and repeated measures. *Nat Methods* 21:83–91

552 Łukasik P, Kolasa MR (2024) With a little help from my friends: the roles of microbial
553 symbionts in insect populations and communities. *Philos Trans R Soc B Biol Sci*
554 379:20230122

555 Łukasik P, Newton JA, Sanders JG, et al (2017) The structured diversity of specialized
556 gut symbionts of the New World army ants. *Mol Ecol* 26:3808–3825

557 McMurdie PJ, Holmes S (2013) phyloseq: an R package for reproducible interactive
558 analysis and graphics of microbiome census data. *PloS One* 8:e61217

559 Meyer JB, Gallien L, Prospero S (2015) Interaction between two invasive organisms on
560 the European chestnut: does the chestnut blight fungus benefit from the
561 presence of the gall wasp? *FEMS Microbiol Ecol* 91:fiv122

562 Mozhaitseva K (2023) Genetic basis of gall formation and modes of reproduction in gall
563 wasps (Hymenoptera, Cynipidae)

564 Murrell P (2002) The grid graphics package. *R News* 2:14–19

565 Mutun S, Dinç S (2023) The role of Anatolia between Asia and Europe: A case study of
566 oak gall wasp species, *Cynips quercus* (Hymenoptera, Cynipidae). *J Adv Res*
567 *Nat Appl Sci* 9:228–236

568 Mutun S, Dinç S, Çimen E (2019) Population genetic structure and phylogeography of
569 the oak gall wasp *Andricus chodjii* (Hymenoptera: Cynipidae) in Turkey as
570 inferred from mitochondrial and nuclear DNA sequences. *Zool Middle East*
571 65:245–255

572 Nowak KH, Hartop E, Prus-Frankowska M, et al (2025) What lurks in the dark? An
573 innovative framework for studying diverse wild insect microbiota. *Microbiome*
574 13:186

575 Pawłowski TA, Staszak AM, Karolewski P, Giertych MJ (2017) Plant development
576 reprogramming by cynipid gall wasp: proteomic analysis. *Acta Physiol Plant*
577 39:114

578 Raman A (2021) Gall-inducing insects and plants. *Curr Sci* 120:66–78

579 Raman A, Wheatley W, Popay A (2012) Endophytic fungus-vascular plant-insect
580 interactions. *Environ Entomol* 41:433–447

581 Ravenscraft A, Coon KL (2026) Transient microbes in insects: fleeting but functional.
582 *Annu Rev Entomol* 71:

583 Rognes T, Flouri T, Nichols B, et al (2016) VSEARCH: a versatile open source tool for
584 metagenomics. *PeerJ* 4:e2584

585 Rohfritsch O (2008) Plants, gall midges, and fungi: a three-component system. *Entomol*
586 *Exp Appl* 128:208–216

587 Rokas A, Atkinson R, Nieves-Aldrey J, et al (2002) The incidence and diversity of
588 *Wolbachia* in gallwasps (Hymenoptera; Cynipidae) on oak. *Mol Ecol* 11:1815–
589 1829

590 Serdo DF, Degaga EG (2023) The hidden world of insect-plant interactions: a review.
591 *SINET-Ethiop J Sci* 46:356–383

592 Siedlecki I, Kochanowski M, Bąk I, et al (2026) Whole forest in a pouch? Methods
593 converge in uncovering wood ants' fungal and bacterial microbiota. *bioRxiv*
594 2026–03

595 Surmacz B, Stec D, Prus-Frankowska M, et al (2024) Pinpointing the microbiota of
596 tardigrades: What is really there? *Environ Microbiol* 26:e16659

597 Takeda S, Hirano T, Ohshima I, Sato MH (2021) Recent progress regarding the
598 molecular aspects of insect gall formation. *Int J Mol Sci* 22:9424

599 Turlousse DM, Yoshiike S, Ohashi A, et al (2017) Synthetic spike-in standards for high-
600 throughput 16S rRNA gene amplicon sequencing. *Nucleic Acids Res* 45:e23–e23

601 Wang Y, Zhang Y, Li R, et al (2025) Host plant dependence of the symbiotic microbiome
602 of the gall-inducing wasp *Trichagalma acutissimae*. *Insects* 16:652

603 Ward AK, Bagley RK, Egan SP, et al (2022) Speciation in Nearctic oak gall wasps is
604 frequently correlated with changes in host plant, host organ, or both. *Evolution*
605 76:1849–1867

606 Washburn G, Van Bael SA (2017) Fungal diversity in galls of baldcypress trees. *Fungal*
607 *Ecol* 29:85–89

608 Wickham H (2016) Data analysis. In: *ggplot2: elegant graphics for data analysis*.
609 Springer, pp 189–201

610 Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: a fast and accurate Illumina
611 Paired-End reAd mergeR. *Bioinformatics* 30:614–620

612 Zhao G-Z, Zhu T-R, Zeng Y, Zhu D-H (2021) *Wolbachia* infection in six species of gall
613 wasps and their parasitoids. *J Asia-Pac Entomol* 24:21–25

614

615