What lurks beneath the humped back? - The patterns of bacterial abundance, diversity and distribution across communities of a dipteran family

Karol H. Nowak[1,2], Emily Hartop[3], Monika Prus-Frankowska[1], Mateusz Buczek[1], Michał Kolasa[1], Piotr Łukasik[1]

1) Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, Kraków, Poland
2) Doctoral School. PhD Programme in Biology, Jagiellonian University, Kraków, Poland
3) Department of Natural History, NTNU University Museum, Norwegian University of Science and Technology, Trondheim, NO-7491, Norway

Abstract

Symbiotic microorganisms have diverse effects on insect life history traits and fitness and can be transmitted within populations and across species. At the level of populations and communities, symbionts can shape insect responses and adaptation to environmental challenges. Despite their importance, our understanding of microbiota in natural insect communities remains limited, particularly for dark taxa - hyperdiverse, abundant, broadly distributed, yet severely understudied groups.

Here, we asked about the abundance, diversity, and distribution of microbiota in natural communities of insects in one such dark taxon, scuttle flies (Diptera: Phoridae), with particular emphasis on their dominant facultative endosymbiont, *Wolbachia*. We did this using high-throughput host and bacterial marker gene amplicons sequencing, including microbiota quantification, for a set of 1,842 humped-back flies representing ca. 186 species from six sites in Northern Sweden.

The resulting dataset, likely the largest microbiota survey for a non-model insect clade to date, provided novel and biologically realistic insights into host-microbe symbioses. We observed significant differences in bacterial abundance among individual insects, with variations up to four orders of magnitude even among flies of the same sex, species, and population. Females consistently harboured more bacteria than males, at least partly due to the consistently higher prevalence and abundance of facultative endosymbionts, especially *Wolbachia* and *Rickettsia*. Within this group, we noted consistent patterns in the prevalence of symbiont genera between host species. Other less abundant bacteria included *Serratia*, *Providencia*, *Carnobacterium*, *Pseudomonas*, *Erwinia*, *Amycolatopsis* and *Spiroplasma*.

Our study highlights the critical role of high-throughput sequencing and integrative methods in revealing the complexity and ecological significance of symbiotic microorganisms in dark taxa, advancing our understanding of their influence on natural communities.

Key words: Insecta, microbiome, endosymbiosis, next-generation sequencing, Illumina, ecological community screening

1. Introduction

1.1 Research background

Symbiotic microorganisms comprise billions of species and strains and include almost all major taxonomic groups of life, such as bacteria, fungi, archaea, viruses, and more. This taxonomic diversity is reflected in their functional diversity - the variety of effects they have on their host, ranging from strict mutualism to pathogenicity (Hammer et al., 2019). Insects are an excellent system for studying microbial symbionts, as they have been associated with a vast taxonomic variety of microorganisms at different points in their evolutionary history, and these microbes can play very different roles in their biology (Perreau & Moran, 2022). Depending on the nature and stability of the association, tissue localisation, transmission type, and genomic features, insect symbionts can be broadly divided into distinct functional categories (Kolasa & Łukasik, 2023). Closed nutritional symbioses involve strictly maternally transmitted, obligatory endosymbionts, essential for host survival, enabling evolutionary adaptations to nutrient-limited diets. Mixed facultative endosymbionts are bacteria that often influence insect reproduction through mechanisms such as male killing or feminisation, cytoplasmic incompatibility and induction of parthenogenesis. They may provide defence against biotic and abiotic stressors and can spread rapidly within and between populations through a mix of vertical and horizontal transmission. Host-adapted ectosymbionts live on the internal or external surfaces of the host, forming variable multi-species communities that may be transmitted vertically, socially, or acquired each generation from the environment. Microbial pathogens primarily transmit horizontally and are generally detrimental to host fitness, although some may use insects as vectors while being harmful only towards plants and vertebrates. Transient microbes are mostly non-stable associates that may temporarily colonise insects, often influencing microbial community profiles without forming consistent infections.

The symbionts' impact on the life history traits of individual insects can be reflected at higher levels of biological organisation - within populations and, likely, multi-species communities. For example, widespread symbionts capable of male-killing or feminisation might influence the sex ratio observed in a population. At the evolutionary scale, symbionts can be viewed as a dynamic reservoir of genetic diversity for their hosts and a source of possible adaptations. Newly acquired symbionts protecting against certain biotic and abiotic stresses can give the individuals harbouring them a selective advantage (B. Zhang et al., 2019). Consequently, the genes of the symbiont-bearing individuals can spread more easily within the population, and the overall species range might broaden (Himler et al., 2011). Since species do not exist in a vacuum, microbiota can indirectly confer beneficial or negative effects on other organisms in the ecosystem. For instance, by protecting from parasitism or predation, they might limit the niche of the species feeding on their hosts.

A particularly well-studied example of an insect symbiont which can significantly impact insect biology at multiple ecological levels comes from the bacterial genus *Wolbachia*. These endosymbionts, sharing ancestors with mitochondria, infect about half of all insect species and

many other arthropods and nematodes (Hilgenboecker et al., 2008; Zug & Hammerstein, 2012). Their widespread prevalence stems from efficient vertical transmission via the female germline and horizontal transfer through behaviours such as cannibalism, predation, parasitism, hybrid introgression, and shared ecological niches (Kaur et al., 2021). Although historically classified mainly as parasitic reproductive manipulators inducing male killing, parthenogenesis or feminisation, their effects cover a broad range between mutualism and parasitism. Various strains can enhance host fitness by providing essential nutrients such as B vitamins (Hosokawa et al., 2010; Moriyama et al., 2015) or granting resistance to viral infections (Teixeira et al., 2008). However, the effect of *Wolbachia* is context-dependent, and by altering their host behaviour and environmental preferences, they might influence the distribution of their host both negatively and positively (Corbin et al., 2017; Truitt et al., 2019). This mechanism might be especially relevant during the ongoing climate change (Lemoine et al., 2020). However, despite the large number of publications on *Wolbachia*, most studies are laboratory-based and neglect their ecological functions and dynamics, as well as their impact on natural insect communities.

1.2 Towards the gap in our knowledge

Despite the importance of microbial symbionts in insect biology, their diversity and distribution are poorly understood. The number of studies investigating the role of insect symbionts and the potential ecological repercussions remains limited outside of non-model systems and laboratory studies. Even in cases where the data is available, there is a strong bias towards model organisms, such as with the article presenting more than a thousand *Wolbachia* genomes, of which 85% assemblies belonged to the strains obtained from just two model *Drosophila* species (Scholz et al., 2020).

There are, however, multiple reasons why studying the composition and significance of the microbiota in wild insect species is challenging. First, the sheer diversity and species richness of their hosts. The most extensive microbial study up to date - the Earth Microbiome Project, has processed more than 27,000 diverse microbial samples (Shaffer et al., 2022). Although the scale of the project was immense, the total number of samples is still smaller than the diversity of insects, even in less biodiverse regions, such as an estimated 33,000 species in Sweden (Ronquist et al., 2020). These thousands of species are often seasonal, live in particular microhabitats, and their presence may depend on many environmental factors, making consistent sampling challenging. Additionally, many of these insect species may represent ‘dark taxa’ - groups that are especially cryptic or difficult to collect and identify, occupy a narrow niche, or may be present in low abundance, further complicating attempts to understand the composition of the communities and their dynamics.

Second, despite recent developments in microbial diversity and distribution research and many broad initiatives, we are still far from understanding the entire scope of their symbiotic relationships. There is a massive gap in our knowledge regarding abundance, taxonomy, as well as spatial and temporal changes in the diversity and distribution of microbes. Genetic reference databases for both hosts and microorganisms are incomplete and contain errors.

The third major problem involves the methodological limitations of generating high-quality data for large numbers of samples and analysing them efficiently. The issues include inconsiderate experimental design, non-representative sample collection, and improper sample storage (Knight et al., 2018), multiple contamination sources present during subsequent workflow steps (Salter et al., 2014), and cross-contamination between samples (Esling et al., 2015). Since this branch of research is relatively new and dynamic, many tools for analysing the data are still being developed and modified. These results in an overwhelming diversity of specific bioinformatic systems and approaches, often difficult to understand. Another issue is the analysis approach, data presentation and interpretation. Insect microbiota abundance is variable, but often low, and frequently, one of the few endosymbiont strains dominates the microbial community. Considering this, one can argue whether 'gold standard' microbial amplicon data analysis workflows, combining rarefaction curves, relative abundance plots for high-level clades, or NMDS/PCA plots, provide biologically meaningful information.

Fortunately, recent technological and conceptual solutions made tackling large numbers of diverse samples while addressing most obstacles plausible. Technological advancements have significantly reduced DNA sequencing costs and a steady increase in sequencing precision. This progress, along with the steady improvement of advanced bioinformatics tools for efficient data analysis, has made high-throughput research more accessible. Additionally, the field has expanded, integrating complementary methods such as microscopy, various -omics techniques, novel experimental approaches and computational modelling.

1.3 Presentation of the phorid study system

Given the importance of studying insect microbiota and the gaps in knowledge, especially in non-model systems, we specifically chose phorid flies (Diptera: Phoridae), also known as scuttle flies, as our focal group. Phorids represent a classic example of a dark taxon: they are hyperdiverse, abundant, and broadly distributed. They thrive in a wide range of environments, exhibiting diverse life strategies and dietary preferences that allow them to occupy numerous ecological niches. Despite their diverse ecological roles and widespread presence, much of their biology, especially their interactions with microbial symbionts, remains unknown. However, novel advances in molecular techniques, particularly high-throughput barcoding, have facilitated further exploration of the system by accelerating the discovery and delimitation of species within this group. Using this approach, over 500 phorid species from Northern Sweden were discovered, including more than a hundred potentially new to science (Hartop, 2021). Existing data from the barcoding efforts, combined with the diversity, availability, simplicity of sampling and ease of laboratory processing of scuttle flies, offer an excellent avenue to illuminate the hidden microbial diversity within this enigmatic group and provide insights relevant to systems well outside of this clade.

1.4 Goals and approaches

The goal of this study was to comprehensively characterise the microbiota composition patterns across scuttle fly communities from Northern Sweden and the diversity and distribution of their dominant microbial symbionts, especially *Wolbachia*. First, we ask about the diversity and distribution patterns of scuttle fly genotypes and species across six sites in Northern Sweden and two seasons. Then, we compare the abundance and composition of phorid microbiota and their distribution patterns across host genotypes, species, sexes, and sampling sites. Finally, we characterise the distribution, diversity, and specificity of the dominant microbe in the dataset, *Wolbachia*. We tackled all three parts by applying quantitative amplicon sequencing of host and bacterial marker genes to 1,842 taxonomically diverse phorid flies sorted from Malaise traps. We analysed the resulting data using a custom bioinformatic workflow emphasising contamination control, microbiome abundance estimation, and tracking microbial clades and sequence variants across host individuals, populations, and species.

2. Materials and methods

2.1. Field collections

The experimental insects were collected using Malaise traps as a part of the Swedish Insect Inventory Project (SIIP) (Karlsson et al., 2020). The collection was carried out at six sites in northern Sweden, separated by varying distances, from less than 50 m to about 390 km. Within each site, a Malaise trap operated during the 2018 vegetation season in two periods: late spring (June to July) and summer (July to August) (Tab. 1). During the collection, bottles containing 95% ethanol were exchanged between every 7 to 31 days due to access challenges. After collection, insects, stored at - 20 °C, were transported to Station Linné, Sweden, where staff taxonomists sorted them into taxonomic bins, ranging from order to family level (depending on the taxon). From the bin corresponding to the family Phoridae, we aimed to randomly select 88 males and the same number of females representing each of the six sites and two seasons, totalling 2,112 insects. However, for some combinations of sex, site and season, we could not obtain the full set of 88 insects, resulting in the final count of 2,105 individuals. As explained further, amplicon data for 1,842 samples (87.5%) have passed our stringent quality criteria and were used for the analyses (Tab. 1).

Tab. 1: Sampling sites and the number of processed phorid males and females passing the quality criteria

						Spring			Summer		
Site ID	Province	City	Site	Latitude	Longitude	Date	Females	Males	Date	Females	Males
SIIP30 (M38)	Norrbotten	Kätkesuando	hay_meadow	68.111431	23.330286	05/06/2018-28/06/2018	50	81	25/07/2018-01/08/2018	84	77
SIIP32 (M32)	Lappland	Abisko	lakeside	68.355154	18.826526	05/06/2018-28/06/2018	79	80	29/07/2018-07/08/2018	82	85
SIIP35 (M35)	Lappland	Abisko, Nuolja	alpine_meadow	68.361011	18.736851	03/06/2018-03/07/2018	70	85	29/07/2018-07/08/2018	82	73
SIIP38 (M38)	Lappland	Ammarnäs	shrubbery	65.957289	16.205235	01/06/2018-02/07/2018	77	74	22/07/2018-17/08/2018	74	69

SIIP43 (M43)	Norrbotten	Kätkesuando	riverbank	68.111319	23.331281	05/06/2018-28/06/2018	67	77	25/07/2018-01/08/2018	69	64
SIIP25 (M25)	Lappland	Voullerim	urban_garden	66.42997	20.62434	01/06/2018-24/06/2018	85	86	22/07/2018-07/08/2018	87	85

2.2. Amplicon library preparation and sequencing

The insects were processed and prepared for sequencing in the Molecular Ecology Laboratory at the Institute of Environmental Sciences of Jagiellonian University, following a protocol comprehensively described in detail on protocols.io (Buczek et al., 2023). First, each insect was placed in a 2 ml tube with 190 µl of the lysis buffer (0.4 M NaCl, 10 mM Tris-HCl, 2 mM EDTA, 2% SDS), 10 µl proteinase-K and different-sized ceramic beads. After two rounds of 30-second homogenisation with a speed of 5 m/s on a homogenizer (Bead Ruptor Elite, Omni International), the samples were incubated for 2 hours in a water bath set at 55 °C. Subsequently, 40 µl aliquot of the homogenate was combined with 2 µl of the solution containing 10,000 copies of plasmids carrying artificial 16S rRNA amplification targets, Ec5502 (Tourlousse et al., 2016) and for subsequent bacterial target quantification, placed in a 96-well PCR plate, and purified on a magnetic plate using 80 µl SPRI beads (Rohland & Reich, 2012). DNA was eluted in 20 µl of TE buffer.

The amplicon libraries were prepared using a two-step procedure. In the first step, two marker regions of interest were simultaneously amplified using QIAGEN Multiplex Master Mix and template-specific primers with variable-length inserts and Illumina adapter stubs. Specifically, for the V4 region of 16S bacterial rRNA, primers 515F: GTGYCAGCMGCCGCGGTAA and 806R: GGACTACNVGGGTWTCTAAT were used (Parada et al., 2016), and for a 458 bp portion of the mitochondrial cytochrome oxidase I (COI) gene, primers BF3: CCHGAYATRGCHTTYCCHCG and BR2: TCDGGRTGNCCRAARAAYCA (Elbrecht et al., 2019). Furthermore, 500 copies of another plasmid carrying an artificial 16S rRNA target, Ec5001, were added to the mixture (Tourlousse et al., 2016). Purified with SPRI beads, the first-step PCR products were used as a template for the second indexing PCR. The details of both PCR steps are described in the abovementioned protocol on protocols.io. To control for contamination from different sources, for each of the 88 biological samples processed, we used eight different controls: negative controls for the DNA extraction (3), first PCR (2), and indexing PCR stages (1), as well as positive controls (2). The libraries were pooled approximately equimolarly, based on band intensity on an agarose gel. The pool was purified twice using Promega beads (ProNex® Size-Selective Purification System), quality checked using BioAnalyzer 2100 (Agilent), combined with pools representing other projects, and submitted to Swedish National Genomics Infrastructure for sequencing on Illumina NovaSeq SPrime flow cell (2x250 bp read mode). The experimental libraries were sequenced in one flow cell.

2.3. Bioinformatic analysis of amplicon data

The amplicon data were processed using the scripts and files available at the GitHub repository: https://github.com/KarolHub/Phorid-Microbiota. Amplicon sequences will be deposited in GenBank upon publication.

2.3.1. Basic read sorting and processing

Initially, we split the demultiplexed reads into batches corresponding to COI and 16S-V4 targets using a custom Python script (Multisplit.py), which divided the files based on primer sequences and trimmed the primers. For the first batch of reads, processed without informative inserts for detecting crosstalk contamination, the Multisplit.py script was used without crosstalk detection. Further batches of reads, however, were processed with informative inserts, reducing the amount of crosstalk. Further steps relied on the custom LSD.py script, a wrapper tool combining several commands executing PEAR (J. Zhang et al., 2014), VSEARCH (Rognes et al., 2016), and USEARCH (Edgar, 2010) programs. Specifically, we merged the files containing forward and reverse sequences based on a high-quality score (phred score >30). Read pairs that were not merged or did not meet the quality standards were discarded. After merging, we obtained the contigs of 400-430 (COI) and 250-260 bp (16S-V4). Both sets were then dereplicated to retain only unique sequences per sample and their occurrence counts. We discarded singletons and ordered the remaining sequences by a descending number of reads. We denoised each library separately to account for erroneous sequences, using the UNOISE algorithm (Antich et al., 2021) from the USEARCH program. This algorithm corrects possible erroneous sequences by incorporating them into more dominant ones based on their ratio.

Afterwards, we combined all the denoised zero-radius operational taxonomic units (zOTUs, ASVs) into one table, with information on how many times each zOTU occurred in each sample. Then, we clustered the zOTUs into operational taxonomic units (OTUs) based on the 97% similarity of the sequences. Finally, we assigned taxonomic labels (blast) to the obtained OTUs using the SINTAX algorithm (Edgar, 2016), which searched for the matching sequences in customised taxonomic datasets: MIDORI version GB 239 (Leray et al., 2018) for the COI data and SILVA (Quast et al., 2012) version 138 SSU for 16S-V4 data. The cutoff for the assignment was 80% similarity of the sequences, above which the sequences were processed by a bootstrapping algorithm to obtain the certainty of a correct taxonomic assignment. Until this point, we processed the COI and 16S-V4 data in parallel using the same pipeline, which produced two tables with zOTU sequences, read counts, classification to OTUs and taxonomic designations.

2.3.2. COI data analysis

The COI data contained a mix of mitochondrial, bacterial, and chimeric sequences, which we separated based on their taxonomy assignment. To obtain barcodes, we discarded sequences shorter or longer than the targeted 418 nucleotides, as well as bacterial, non-arthropod, unassigned and chimeric sequences. Next, we assessed the number of reads in the most abundant OTU classified as an arthropod and calculated its ratio relative to all Arthropoda OTUs combined. We excluded libraries from further analyses if they had fewer than 300 reads of the dominant OTU or if the aforementioned read ratio was less than 95% - giving a total of 1845 filtered libraries. We also eliminated 3 samples based on the lack of spike-in reads in the 16S rRNA data (see section below). Then, we designated the most abundant sequence within the dominant OTU of each sample as the individual's barcode, dereplicated these barcodes, and clustered them with sequences from earlier barcoding studies (Hartop, 2021) to generate a dendrogram. Given the complex phylogeny of phorid flies, we manually reviewed the proximity of each barcode within the dendrogram to determine its match with existing morphospecies. Species names were assigned according to BLAST results, retaining only labels with a match of 97% or higher, leading to varying uncertainty at the species or genus level.

The bacterial COI data included sequences assigned to the order Rickettsiales, which we used in further analyses. We filtered out non-*Wolbachia* sequences, chimeric sequences, and those of length other than 418.

2.3.3. 16S rRNA amplicon data filtering and abundance estimation

We decontaminated the 16S-V4 data from 1,845 samples remaining after COI-level filtering using a custom QUACK.py script. This script removed zOTUs whose relative abundance in any negative samples was at least five times greater than in insect samples, together with sequences classified as Archaea, Eukaryota, chloroplasts, mitochondria and chimaeras. After calculating the ratio between the spike-in and remaining symbiont reads in each sample, the spike-in sequences were also removed. Afterwards, we calculated the absolute abundance of bacteria for each sample by multiplying the symbiont to extraction spike-ins ratio, the number of inserted extraction spike-in copies (10,000), and the inverse proportion of insect homogenate used during the DNA extraction (5). Three samples without extraction spike-in reads were discarded, along with control samples, resulting in a final dataset of 1,842 samples.

3. Results

3.1. COI amplicon-based insights into the diversity and distribution of targetted insects

3.1.1. Data summary

After the filtering described in the method section, we kept an average of 77 individuals for each of the 24 site-season-sex categories, including 936 males and 906 females - 1,842 in total (Tab. S1). The supplementary material is available in the GitHub repository https://github.com/KarolHub/Phorid-Microbiota.

For individuals passing the filtering criteria, we obtained, on average, 27,130 COI reads (min 373, max 191,926, median 20,422) (Tab. S2). On average, the most abundant insect zOTU, the barcode, among samples accounted for 98.8% of the reads (min 52.2%, max 100%, median 99.5%). We designated this dominant zOTU as the individual's barcode. The COI barcodes of 1,842 phorid individuals represented 480 genotypes that, through similarity analysis and comparisons with previously characterised material (Hartop, 2021), we grouped into 186 species (Tab. S3). The 359 of our 480 reconstructed COI-based genotypes were confidently matched (<3% sequence difference) to established morphospecies through a phylogenetic comparison against previously curated sequences, and 242 matched them perfectly. Although 142 genotypes obtained a full species name, 338 of the reconstructed COI-based genotypes belonged to clades where species delimitation proved challenging, resulting in classification uncertainty. As we were not confident about the identification, we decided to classify them at the genus or tribe level only. For simplicity, in instances where the phorid taxonomy is not essential, we numbered all species by decreasing abundance and will refer to them by these numbers. Within these species, the proportion of captured males and females varied. Additionally, in 94 species with fewer than 13 individuals, only one sex was captured.

3.1.2. Species counts and distributions

The most abundant 20 species accounted for 59.4% of the 1842 processed individuals, with considerable variation in species distribution across the adjacent sampling dates. As expected, differences were present in the numbers of specimens classified into different species. The most abundant sp.1 (Megaselia sp. 001) comprised 227 individuals. On the other end of the spectrum, 54 species were represented by a single individual. Most species, including the most abundant Megaselia sp. 001, were abundant during only one of the seasons and represented by few or no specimens during the other season.

Within the same season across all sites, we found an overlapping set of abundant species, with the top five most abundant ones found at every site. At the same time, even the most distant locations (M38 and M43) shared the same dominant species. However, none of the abundant species were site-specific, as all species represented by at least 15 individuals were present in more than one site. Notably, the M25 site, located in a warmer, suburban area and at a lower altitude, had a distinct community composition, with 33.8% of its individuals representing species absent from the remaining sites. However, 14 of the 20 most abundant species were present in the M25 location (Fig. 1A, B). Principal Coordinates Analysis (Fig. 1C) further emphasised the dissimilarity of species composition between the seasons. At the same time, it highlighted an overlap across the sites of comparable climate conditions (M35, M43, M32, M28, M38), especially during the summer, while pointing toward the distinct composition of the M25 site community.

3.1.3. Within species genetic diversity

Although most collected species belonged to the genus *Megaselia*, we observed a substantial COI genotypic variation within the species represented by multiple individuals (Fig. 1D). Within each of the five most abundant species in the collection, there were between 10 and 33 genotypes. While almost all species harboured multiple genotypes, most were present at low abundance, with only one or two genotypes dominating within a species. Within the top five most abundant species, 59 of the 85 genotypes were represented by a single individual. Many low-abundance genotypes often diverged from the dominant genotype by a single nucleotide. In cases where multiple abundant genotypes were present, we observed a variation in genotype proportions across sites, although there is little evidence of rigid geographic structuring within the same species (Fig. 1E).

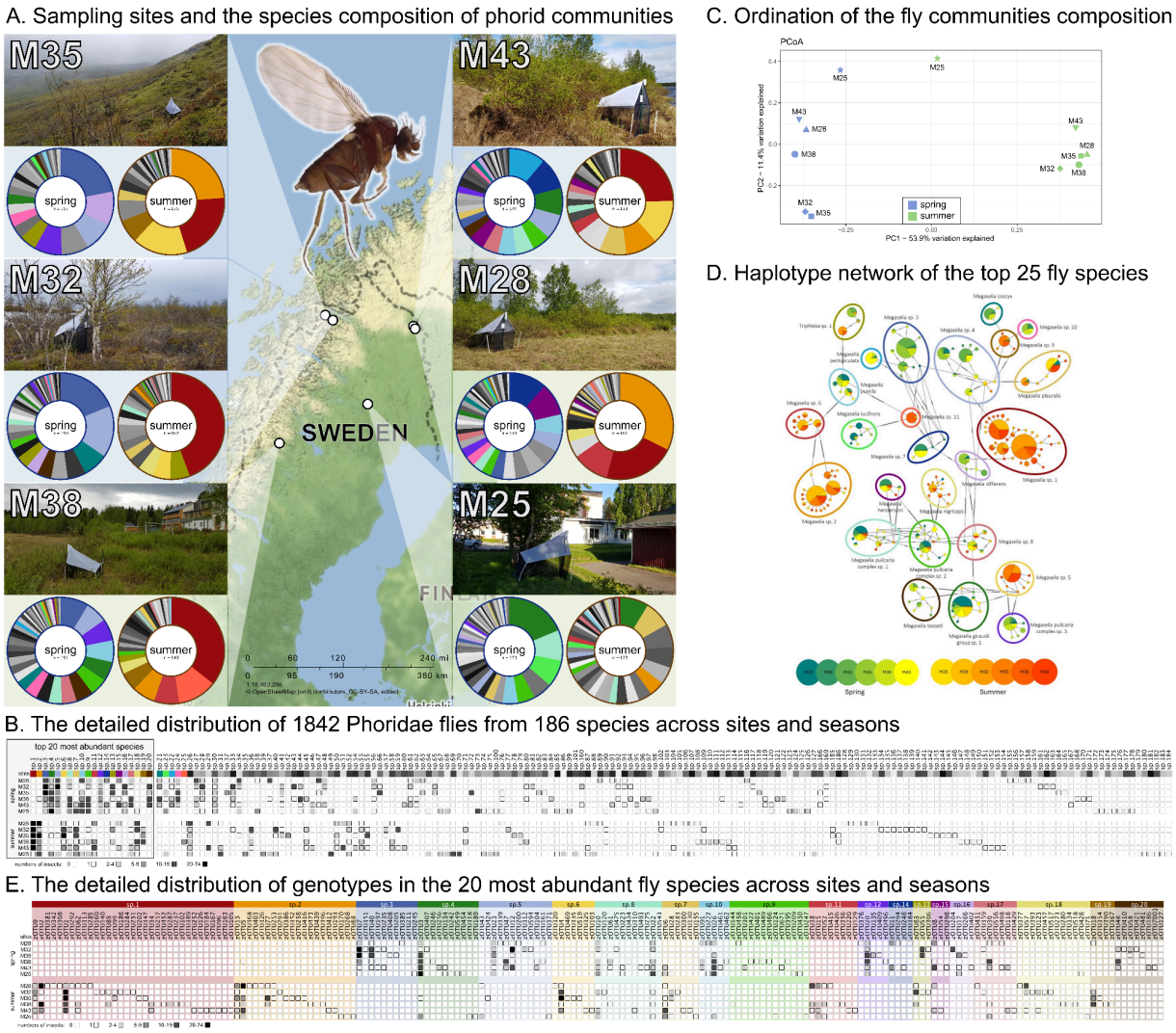

Fig. 1: Distribution and diversity of phorid flies across two seasons and six sites in Northern Sweden. A. The map of the sampling sites and the sampled community composition across sites and dates. B. The diversity and distribution of sampled species across sites and dates. C. Genotype-level diversity within selected fly species and their distribution among sites. D. Comparison of the sampled phorid collection sites based on species composition ordinated using Principal Coordinates Analysis (PCoA) with Bray-Curtis dissimilarity. E. Haplotype network of the genotypes of the most abundant 25 species, coloured based on the site and season.

3.2. Abundance, diversity and distribution of phorid microbiota (abundance and composition go in tandem)

3.2.1. Data summary

Across all the specimens, we obtained a substantial, although varying, number of 16S-V4 reads (min 2, max 295,672, median 20,222). After filtering and decontamination, the samples contained between 29 and 227,059 16S-V4 reads (median 5,264). In total, there were 91,897 unique 16S-V4 zOTUs in all the insect samples before the decontamination, from which 56,132 remained after the filtering (Tab. S4). Decontaminated bacterial zOTUs were then clustered into 9,589 OTUs based on a 97% similarity of sequences (Tab. S5). 43.7% of total decontaminated bacterial reads belonged to various *Wolbachia* genotypes, while other bacteria were represented by a gradually decreasing share of reads, including *Rickettsia* - 9.9%, *Serratia* - 9.0%, *Providencia* - 5.3% and *Spiroplasma* - 4.2%.

3.2.2. Microbiota abundance, composition and prevalence

We observed high differences in the symbiont-to-spike-in ratio across the samples, ranging between 0.015 and 354.8 (median 1.37) (Tab. S6). These uneven proportions suggest several orders of magnitude of variation in microbiome abundance among phorid specimens. The calculated absolute abundance of bacterial gene copies ranged between 765 and 24,008,154 (median 89,360). These differences were visible between individuals from the same species and sexes, with frequently observed four orders of magnitude differences (Fig. 2A). The abundance was consistently higher in females, often harbouring about ten times more microbes than males of the same species.

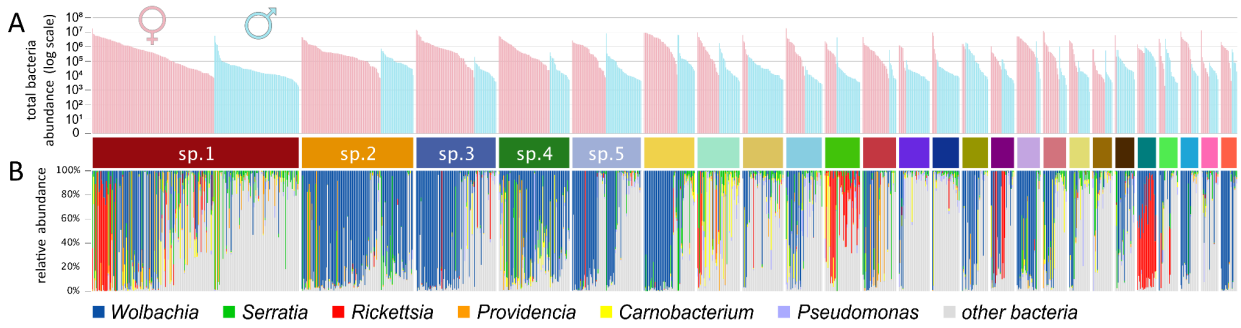

Fig. 2: Distribution of absolute abundance and microbiota composition across 1,193 samples of the 25 dominant phorid species. A.16S-V4 rRNA copy number distribution across individual insects, estimated based on the stable number of 10,000 copies of quantification spike-in added during the DNA extraction step. Note the logarithmic scale. B. The distribution of relative abundances of the six most abundant bacterial OTUs out of 9,589. Category 'other bacteria' contains all the remaining bacterial OTUs combined.

Several symbiont genera appeared to dominate the microbiota composition (Tab. S7), with facultative endosymbiont species being the most prominent. Specifically, *Wolbachia* comprised over 30.3% of bacterial relative abundance. Despite the presence of 135 16S-V4 *Wolbachia* OTUs, a single dominant OTU accounted for more than 99.5% of the total *Wolbachia* abundance. Subsequent prevalent symbiont OTUs, including *Serratia*, *Rickettsia*, *Providencia*, *Carnobacterium,* and *Pseudomonas,*

comprised 5.4%, 5.1%, 3.0%, 2.9%, and 2.7% of the relative abundance, respectively. We noted a striking link between microbial abundance and the presence of dominant microbes (Fig. 2B). Specifically, when looking at microbial profiles of males and females sorted by the decreasing abundance, we observed a clear connection between the sex and a significantly increased total absolute abundance of bacterial gene copies. These differences are largely due to the consistently higher prevalence and abundance of facultative endosymbionts, especially *Wolbachia* and *Rickettsia*, in females rather than males. However, we noted variance in the harboured endosymbionts taxa between species and sexes. For example, within sp.1, *Rickettsia* is mostly present in females but almost absent in males. A considerable level of variation is visible at the level of population as well. The same species of phorid flies presented different patterns of microbiota abundance and composition across different populations. Within the most common sp.1, high bacterial abundance in males was observed primarily in individuals from the M38 site (Fig. 3A). At the same time, *Rickettsia* was detected in high prevalence in three out of five sufficiently numerous populations (Fig. 3B).

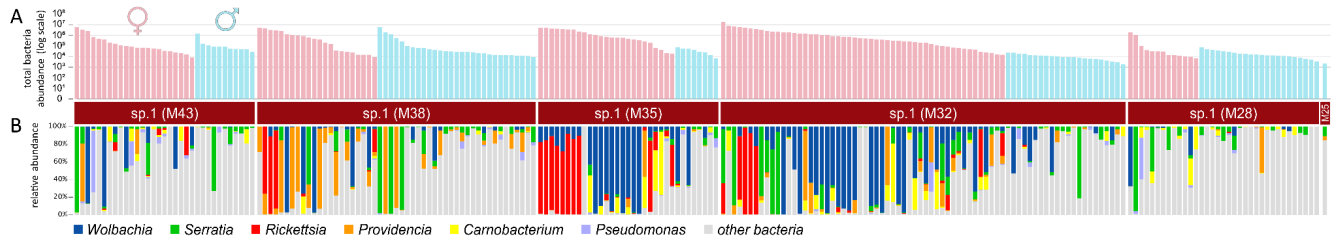

Fig. 3: Distribution of bacterial absolute abundance and distribution across the six sites within the most abundant species (Megaselia sp. 001). A.16S-V4 rRNA copy number distribution across individual insects, note the logarithmic scale. B. Distribution of the relative abundance of the six most abundant bacterial OTUs. Category 'other bacteria' contains all the remaining bacterial OTUs combined.

3.3. Wolbachia diversity and distribution

The COI marker region reads used for barcoding the phorid individuals contained various sequences beyond the intended barcode. Of these, 99.5% were classified as insect sequences, with 98.8% matching the barcode and 0.7% corresponding to other host-derived, environmental, contaminant, or low-quality reads. The remaining 0.5% of sequences predominantly belonged to *Wolbachia*. About 47% of samples contained, on average, 269 reads per sample (min 2, max 4,875, median 100). From these, we obtained 258 unique zOTUs grouped into 23 OTUs. A comparison of the 16S-V4 and COI data revealed about 95% overlap in *Wolbachia* presence across the samples: about 25% of infected males and 75% of infected females. However, the COI data offered much higher resolution at the genotype level, allowing for a detailed analysis of *Wolbachia* diversity across phorid species. The distribution of *Wolbachia* variants was non-random, with most species dominated by a single variant (Fig. 4). Interestingly, the same *Wolbachia* variant was found in different species of flies. Additionally, some species, such as sp.5, harboured more than one dominant *Wolbachia* variant, with consistent double infections observed in females.

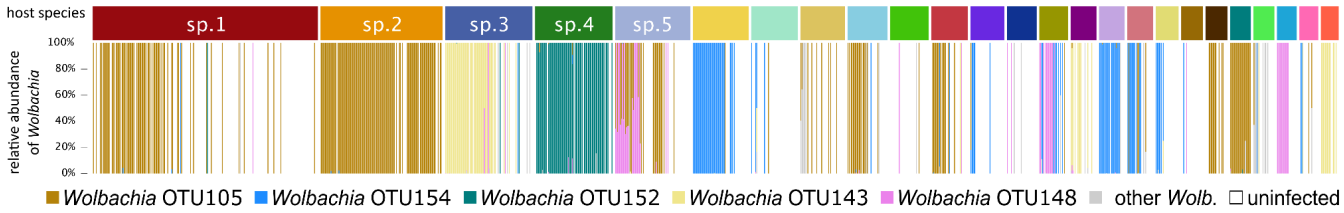

Fig. 4: Distribution of the relative shares of *Wolbachia* OTUs obtained from the COI data of the most common 25 phorid species.

4. Discussion

4.1 The diversity and distribution of targeted insects

By combining preliminary taxonomical categorisation based on morphology with a custom bioinformatic procedure, we efficiently delimited phorid species and analysed their spatiotemporal distribution. The species composition changed significantly between adjacent seasons, suggesting rapid species turnover, likely influenced by their phenology. The scale of this turnover and the variation in species composition at the suburban M25 site indicate considerable species richness within Phoridae. Both of these observations are consistent with the previous studies (Hartop, 2021). Additionally, we report a major overlap in the composition of dominant species among habitats, implying that some of the most abundant insect species near the Arctic Circle may have a broad geographical distribution. However, this pattern might be influenced by the selectivity of the Malaise traps (Disney, 1994). To address this issue in the future, it would be beneficial to complement Malaise traps with emergence traps, light traps, or other sampling methods to improve species coverage (Borkent et al., 2018). Finally, we present how the barcoding data provides a robust framework for understanding insect microbiota patterns.

4.2 Abundance, diversity and distribution of phorid microbiota

Our screening of individual insects revealed a significant variation in bacterial abundance, often spanning several orders of magnitude within a single fly species, sex, and site. While factors such as insect age, diet, or interactions with other organisms might contribute to this variation (Li et al., 2023), the size of the insects is unlikely to be a major factor, as 98% of the flies were classified into the same size category. Despite the substantial individual differences, a clear pattern emerged: females consistently showed higher bacterial abundance than males, likely due to the dominance of facultative endosymbionts in their microbiota (Fig. 2A, B). These endosymbionts, often present in large numbers, can occupy internal tissues, which, compared to external surfaces, offer ample space for reproduction and are protected against environmental threats (Perlmutter & Bordenstein, 2020). Additionally, endosymbionts may further influence the host's internal environment by modulating the immune system and other physiological parameters to influence other microorganisms (Duan et al., 2020; Eleftherianos et al., 2013), which could explain the higher symbiont titre observed in infected individuals. Aside from dominant endosymbionts, the flies also harboured various other bacteria, though most were significantly less abundant. Nevertheless, these less prevalent bacteria might play crucial roles. For instance, *Providencia*, *Serratia*, *Erwinia*, and *Enterobacter* are significant symbionts in both Hymenoptera and

Diptera (Noman et al., 2020; Wang & Brucker, 2019). Bacteria such as *Pseudomonas*, known for their role in bark beetles, can provide nutrients and protection against tree defences (Saati-Santamaría et al., 2021), although some strains, especially found in solitary bees can also be regarded as environmental contamination or pathogens (Fernandez De Landa et al., 2023). Additionally, we observed a consistent pattern of higher bacterial diversity in males than females. A similar phenomenon has been described before in Hymenoptera studies (Guo et al., 2020), linked to differences in immune capabilities between the sexes and the endosymbionts like *Rickettsia* changing the microbiota composition by influencing the immune response of their host. The high prevalence of endosymbionts among the studied insects is likely due to the endosymbionts' reliable vertical transmission capabilities, allowing them to spread efficiently within populations and form stable symbiotic or mutualistic relationships, which can enhance host fitness (McLean et al., 2018). The disproportion in the prevalence of these bacteria between sexes can be attributed to the bacteria's ability to manipulate insect reproduction through mechanisms such as feminisation and male killing, which reduce the proportion of infected males (Kageyama et al., 2012; Werren et al., 2008). A different kind of manipulation, cytoplasmic incompatibility, can increase the prevalence of infected individuals in the population by reducing the reproductive success of the uninfected females (Engelstädter & Telschow, 2009). Reproductive manipulations and benefits provided by the endosymbionts have facilitated symbiont sweeps—a rapid spread of beneficial bacteria in the host population. This phenomenon has been observed with the spread of *Wolbachia*, which increased host fecundity (Kriesner et al., 2013), *Rickettsia*, which enhanced host growth and reproduction (Himler et al., 2011), and *Spiroplasma*, which protected against the sterilising effects of nematodes (Jaenike et al., 2010).

4.3 Distribution, diversity and specificity of Wolbachia

Wolbachia's dominance in the microbiota composition is unsurprising, as these bacteria are among the most successful endosymbionts globally. Their success can be attributed to effective transmission within and sometimes across host species, manipulation of host sex and reproduction, and possible provision of protective or nutritional benefits (Sanaei et al., 2021). Our data reveal strong host specificity of *Wolbachia*, although a few *Wolbachia* variants were observed dominating several species. A parallel study in *Drosophila* demonstrated a rapid spread of *Wolbachia* across multiple species within 30,000 years (Turelli et al., 2018). To determine if similar processes happened in scuttle flies, further metagenomic studies are needed to explore phylogenetic relationships and transmission pathways.

Conclusion

Symbiotic microorganisms have the potential to alter their host's response to environmental stressors, thereby influencing their niche and adaptability. Considering the ecological importance of insects, it is necessary to intensify the efforts to fill the gaps regarding the dynamics of both insect communities and their symbionts (Lemoine et al., 2020). Thus, we imply a need for more broad studies of microbiota dynamics in insect communities using high-throughput technologies and approaches, possibly similar to the one presented in this study.

Acknowledgements

The project was supported by the Polish National Agency for Academic Exchange grant PPN/PPO/2018/1/00015 and Polish National Science Centre grants 2018/31/B/NZ8/01158 and 2021/41/B/NZ8/04526 (to P.Ł.)

References

Antich, A., Palacin, C., Wangensteen, O. S., & Turon, X. (2021). To denoise or to cluster, that is not the question: Optimizing pipelines for COI metabarcoding and metaphylogeography. *BMC Bioinformatics*, *22*(1), 177. https://doi.org/10.1186/s12859-021-04115-6

Borkent, A., Brown, B. V., Adler, P. H., Amorim, D. D. S., Barber, K., Bickel, D., Boucher, S., Brooks, S. E., Burger, J., Burington, Z. L., Capellari, R. S., Costa, D. N. R., Cumming, J. M., Curler, G., Dick, C. W., Epler, J. H., Fisher, E., Gaimari, S. D., Gelhaus, J., … Zumbado, M. A. (2018). Remarkable fly (Diptera) diversity in a patch of Costa Rican cloud forest: Why inventory is a vital science. *Zootaxa*, *4402*(1), 53. https://doi.org/10.11646/zootaxa.4402.1.3

Buczek, M., Prus-Frankowska, M., & Łukasik, P. (2023). *Quantitative multi-target amplicon sequencing workflow v1*. https://doi.org/10.17504/protocols.io.36wgq351ylk5/v1

Corbin, C., Heyworth, E. R., Ferrari, J., & Hurst, G. D. D. (2017). Heritable symbionts in a world of varying temperature. *Heredity*, *118*(1), 10–20. https://doi.org/10.1038/hdy.2016.71

Disney, R. H. L. (1994). *Scuttle Flies: The Phoridae*. Springer Netherlands. https://doi.org/10.1007/978-94-011-1288-8

Duan, X.-Z., Sun, J.-T., Wang, L.-T., Shu, X.-H., Guo, Y., Keiichiro, M., Zhu, Y.-X., Bing, X.-L., Hoffmann, A. A., & Hong, X.-Y. (2020). Recent infection by Wolbachia alters microbial communities in wild Laodelphax striatellus populations. *Microbiome*, *8*(1), 104. https://doi.org/10.1186/s40168-020-00878-x

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Edgar, R. C. (2016). *SINTAX: A simple non-Bayesian taxonomy classifier for 16S and ITS sequences*.

https://doi.org/10.1101/074161

Elbrecht, V., Braukmann, T. W. A., Ivanova, N. V., Prosser, S. W. J., Hajibabaei, M., Wright, M., Zakharov, E. V., Hebert, P. D. N., & Steinke, D. (2019). Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, *7*, e7745. https://doi.org/10.7717/peerj.7745

Eleftherianos, I., Atri, J., Accetta, J., & Castillo, J. C. (2013). Endosymbiotic bacteria in insects: Guardians of the immune system? *Frontiers in Physiology*, *4*. https://doi.org/10.3389/fphys.2013.00046

Engelstädter, J., & Telschow, A. (2009). Cytoplasmic incompatibility and host population structure. *Heredity*, *103*(3), 196–207. https://doi.org/10.1038/hdy.2009.53

Esling, P., Lejzerowicz, F., & Pawlowski, J. (2015). Accurate multiplexing and filtering for high-throughput amplicon-sequencing. *Nucleic Acids Research*, *43*(5), 2513–2524. https://doi.org/10.1093/nar/gkv107

Fernandez De Landa, G., Alberoni, D., Baffoni, L., Fernandez De Landa, M., Revainera, P. D., Porrini, L. P., Brasesco, C., Quintana, S., Zumpano, F., Eguaras, M. J., Maggi, M. D., & Di Gioia, D. (2023). The gut microbiome of solitary bees is mainly affected by pathogen assemblage and partially by land use. *Environmental Microbiome*, *18*(1), 38. https://doi.org/10.1186/s40793-023-00494-w

Guo, C., Peng, X., Zheng, X., Wang, X., Wang, R., Huang, Z., & Yang, Z. (2020). Comparison of bacterial diversity and abundance between sexes of *Leptocybe invasa* Fisher & La Salle (Hymenoptera: Eulophidae) from China. *PeerJ*, *8*, e8411. https://doi.org/10.7717/peerj.8411

Hammer, T. J., Sanders, J. G., & Fierer, N. (2019). Not all animals need a microbiome. *FEMS Microbiology Letters*, *366*(10), fnz117. https://doi.org/10.1093/femsle/fnz117

Hartop, E. (2021). *A multi-faceted approach to a "dark taxon": The hyperdiverse and poorly known scuttle flies (Diptera: Phoridae)*. http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-192276

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., & Werren, J. H. (2008). How many species are infected with Wolbachia? – A statistical analysis of current data: Wolbachia infection rates. *FEMS Microbiology Letters*, *281*(2), 215–220.

https://doi.org/10.1111/j.1574-6968.2008.01110.x

Himler, A. G., Adachi-Hagimori, T., Bergen, J. E., Kozuch, A., Kelly, S. E., Tabashnik, B. E., Chiel, E., Duckworth, V. E., Dennehy, T. J., Zchori-Fein, E., & Hunter, M. S. (2011). Rapid Spread of a Bacterial Symbiont in an Invasive Whitefly Is Driven by Fitness Benefits and Female Bias. *Science*, *332*(6026), 254–256. https://doi.org/10.1126/science.1199410

Hosokawa, T., Koga, R., Kikuchi, Y., Meng, X.-Y., & Fukatsu, T. (2010). *Wolbachia* as a bacteriocyte-associated nutritional mutualist. *Proceedings of the National Academy of Sciences*, *107*(2), 769–774. https://doi.org/10.1073/pnas.0911476107

Jaenike, J., Unckless, R., Cockburn, S. N., Boelio, L. M., & Perlman, S. J. (2010). Adaptation via Symbiosis: Recent Spread of a *Drosophila* Defensive Symbiont. *Science*, *329*(5988), 212–215. https://doi.org/10.1126/science.1188235

Kageyama, D., Narita, S., & Watanabe, M. (2012). Insect Sex Determination Manipulated by Their Endosymbionts: Incidences, Mechanisms and Implications. *Insects*, *3*(1), 161–199. https://doi.org/10.3390/insects3010161

Karlsson, D., Hartop, E., Forshage, M., Jaschhof, M., & Ronquist, F. (2020). The Swedish Malaise Trap Project: A 15 Year Retrospective on a Countrywide Insect Inventory. *Biodiversity Data Journal*, *8*, e47255. https://doi.org/10.3897/BDJ.8.e47255

Kaur, R., Shropshire, J. D., Cross, K. L., Leigh, B., Mansueto, A. J., Stewart, V., Bordenstein, S. R., & Bordenstein, S. R. (2021). Living in the endosymbiotic world of Wolbachia: A centennial review. *Cell Host & Microbe*, *29*(6), 879–893. https://doi.org/10.1016/j.chom.2021.03.006

Knight, R., Vrbanac, A., Taylor, B. C., Aksenov, A., Callewaert, C., Debelius, J., Gonzalez, A., Kosciolek, T., McCall, L.-I., McDonald, D., Melnik, A. V., Morton, J. T., Navas, J., Quinn, R. A., Sanders, J. G., Swafford, A. D., Thompson, L. R., Tripathi, A., Xu, Z. Z., … Dorrestein, P. C. (2018). Best practices for analysing microbiomes. *Nature Reviews Microbiology*, *16*(7), 410–422. https://doi.org/10.1038/s41579-018-0029-9

Kolasa, M., & Łukasik, P. (2023). *With a little help from my friends: The roles of microbial symbionts in*

insect populations and communities. https://doi.org/10.32942/X2BS45

Kriesner, P., Hoffmann, A. A., Lee, S. F., Turelli, M., & Weeks, A. R. (2013). Rapid Sequential Spread of Two Wolbachia Variants in Drosophila simulans. *PLoS Pathogens*, *9*(9), e1003607. https://doi.org/10.1371/journal.ppat.1003607

Lemoine, M. M., Engl, T., & Kaltenpoth, M. (2020). Microbial symbionts expanding or constraining abiotic niche space in insects. *Current Opinion in Insect Science*, *39*, 14–20. https://doi.org/10.1016/j.cois.2020.01.003

Leray, M., Ho, S.-L., Lin, I.-J., & Machida, R. J. (2018). MIDORI server: A webserver for taxonomic assignment of unknown metazoan mitochondrial-encoded sequences using a curated database. *Bioinformatics*, *34*(21), 3753–3754. https://doi.org/10.1093/bioinformatics/bty454

Li, Y., Chang, L., Xu, K., Zhang, S., Gao, F., & Fan, Y. (2023). Research Progresses on the Function and Detection Methods of Insect Gut Microbes. *Microorganisms*, *11*(5), 1208. https://doi.org/10.3390/microorganisms11051208

Moriyama, M., Nikoh, N., Hosokawa, T., & Fukatsu, T. (2015). Riboflavin Provisioning Underlies *Wolbachia* 's Fitness Contribution to Its Insect Host. *mBio*, *6*(6), e01732-15. https://doi.org/10.1128/mBio.01732-15

Noman, M. S., Liu, L., Bai, Z., & Li, Z. (2020). Tephritidae bacterial symbionts: Potentials for pest management. *Bulletin of Entomological Research*, *110*(1), 1–14. https://doi.org/10.1017/S0007485319000403

Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, *18*(5), 1403–1414. https://doi.org/10.1111/1462-2920.13023

Perlmutter, J. I., & Bordenstein, S. R. (2020). Microorganisms in the reproductive tissues of arthropods. *Nature Reviews Microbiology*, *18*(2), Article 2. https://doi.org/10.1038/s41579-019-0309-z

Perreau, J., & Moran, N. A. (2022). Genetic innovations in animal–microbe symbioses. *Nature Reviews*

Genetics, *23*(1), 23–39. https://doi.org/10.1038/s41576-021-00395-z

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2012). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, *41*(D1), D590–D596. https://doi.org/10.1093/nar/gks1219

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, *4*, e2584. https://doi.org/10.7717/peerj.2584

Rohland, N., & Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, *22*(5), 939–946. https://doi.org/10.1101/gr.128124.111

Ronquist, F., Forshage, M., Häggqvist, S., Karlsson, D., Hovmöller, R., Bergsten, J., Holston, K., Britton, T., Abenius, J., Andersson, B., Buhl, P. N., Coulianos, C.-C., Fjellberg, A., Gertsson, C.-A., Hellqvist, S., Jaschhof, M., Kjærandsen, J., Klopfstein, S., Kobro, S., … Gärdenfors, U. (2020). Completing Linnaeus's inventory of the Swedish insect fauna: Only 5,000 species left? *PLOS ONE*, *15*(3), e0228561. https://doi.org/10.1371/journal.pone.0228561

Saati-Santamaría, Z., Rivas, R., Kolařik, M., & García-Fraile, P. (2021). A New Perspective of Pseudomonas—Host Interactions: Distribution and Potential Ecological Functions of the Genus Pseudomonas within the Bark Beetle Holobiont. *Biology*, *10*(2), 164. https://doi.org/10.3390/biology10020164

Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., Turner, P., Parkhill, J., Loman, N. J., & Walker, A. W. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biology*, *12*(1), 87. https://doi.org/10.1186/s12915-014-0087-z

Sanaei, E., Charlat, S., & Engelstädter, J. (2021). Wolbachia host shifts: Routes, mechanisms, constraints and evolutionary consequences. *Biological Reviews*, *96*(2), 433–453. https://doi.org/10.1111/brv.12663

Scholz, M., Albanese, D., Tuohy, K., Donati, C., Segata, N., & Rota-Stabelli, O. (2020). Large scale

genome reconstructions illuminate Wolbachia evolution. *Nature Communications*, *11*(1), 5235. https://doi.org/10.1038/s41467-020-19016-0

Shaffer, J. P., Nothias, L.-F., Thompson, L. R., Sanders, J. G., Salido, R. A., Couvillion, S. P., Brejnrod, A. D., Lejzerowicz, F., Haiminen, N., Huang, S., Lutz, H. L., Zhu, Q., Martino, C., Morton, J. T., Karthikeyan, S., Nothias-Esposito, M., Dührkop, K., Böcker, S., Kim, H. W., … Zhang, S. (2022). Standardized multi-omics of Earth's microbiomes reveals microbial and metabolite diversity. *Nature Microbiology*, *7*(12), 2128–2150. https://doi.org/10.1038/s41564-022-01266-x

Teixeira, L., Ferreira, Á., & Ashburner, M. (2008). The Bacterial Symbiont Wolbachia Induces Resistance to RNA Viral Infections in Drosophila melanogaster. *PLoS Biology*, *6*(12), e1000002. https://doi.org/10.1371/journal.pbio.1000002

Tourlousse, D. M., Yoshiike, S., Ohashi, A., Matsukura, S., Noda, N., & Sekiguchi, Y. (2016). Synthetic spike-in standards for high-throughput 16S rRNA gene amplicon sequencing. *Nucleic Acids Research*, gkw984. https://doi.org/10.1093/nar/gkw984

Truitt, A. M., Kapun, M., Kaur, R., & Miller, W. J. (2019). *Wolbachia* modifies thermal preference in *Drosophila melanogaster* . *Environmental Microbiology*, *21*(9), 3259–3268. https://doi.org/10.1111/1462-2920.14347

Turelli, M., Cooper, B. S., Richardson, K. M., Ginsberg, P. S., Peckenpaugh, B., Antelope, C. X., Kim, K. J., May, M. R., Abrieux, A., Wilson, D. A., Bronski, M. J., Moore, B. R., Gao, J.-J., Eisen, M. B., Chiu, J. C., Conner, W. R., & Hoffmann, A. A. (2018). Rapid Global Spread of wRi-like Wolbachia across Multiple Drosophila. *Current Biology*, *28*(6), 963-971.e8. https://doi.org/10.1016/j.cub.2018.02.015

Wang, G.-H., & Brucker, R. M. (2019). Genome Sequence of Providencia rettgeri NVIT03, Isolated from Nasonia vitripennis. *Microbiology Resource Announcements*, *8*(3), e01157-18. https://doi.org/10.1128/MRA.01157-18

Werren, J. H., Baldo, L., & Clark, M. E. (2008). Wolbachia: Master manipulators of invertebrate biology. *Nature Reviews Microbiology*, *6*(10), 741–751. https://doi.org/10.1038/nrmicro1969

Zhang, B., Leonard, S. P., Li, Y., & Moran, N. A. (2019). Obligate bacterial endosymbionts limit thermal tolerance of insect host species. *Proceedings of the National Academy of Sciences*, *116*(49), 24712–24718. https://doi.org/10.1073/pnas.1915307116

Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2014). PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, *30*(5), 614–620. https://doi.org/10.1093/bioinformatics/btt593

Zug, R., & Hammerstein, P. (2012). Still a Host of Hosts for Wolbachia: Analysis of Recent Data Suggests That 40% of Terrestrial Arthropod Species Are Infected. *PLoS ONE*, *7*(6), e38544. https://doi.org/10.1371/journal.pone.0038544