Cold and lonely: low-abundance microbiota and no core microbiome across a broad geographic sampling of a mosquito species in Greenland

Diana Rojas-Guerrero[1], Michał Kolasa[1], Mateusz Buczek[1], Monika Prus-Frankowska[1], Karol H. Nowak[1], Tomas Roslin[2], Piotr Łukasik[1]

1. Institute of Environmental Sciences, Faculty of Biology, Jagiellonian University, Kraków, Poland
2. Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden

Abstract

Microbial symbionts play a crucial role in the biology of their insect hosts. In recent years, the mosquito microbiome has emerged as a promising tool for controlling mosquito populations and altering their capacity to transmit diseases. However, while mosquitoes are widespread, most microbiome research focuses on a few model species, leaving many others understudied. In this study, we investigated how temporal and spatial variations influence the microbial abundance and diversity associated with *Ochlerotatus* mosquitoes in Greenland. We did this by conducting high-throughput sequencing of insect and bacterial marker gene amplicons and quantifying bacterial abundance across a large collection of mosquitoes sampled systematically from four regions of Greenland between 2009 and 2023.

Analysis of COI data identified two species within our collection: the widely distributed *O. nigripes* and the less common *O. impiger*. Quantitative 16S rRNA amplicon sequencing revealed a highly variable number of bacterial rRNA gene copies per mosquito, with overall microbiota abundance typically lower (Avg = ~1.2×10^5, Mdn = 1.04×10^4) than that estimated for both lab-reared *Aedes aegypti* and wild *Culex pipiens* from Poland. Despite identifying diverse bacterial taxa, including *Pseudomonas*, *Wolbachia*, and *Bartonella*, we observed significant microbiome variability between individual mosquitoes. Notably, there were no consistent patterns of microbiota distribution across sexes, geographic locations, or sampling years. These findings suggest that *Ochlerotatus* mosquitoes from the surveyed populations in Greenland may not rely on abundant or specialized microbiota for their reproduction or nutrition, indicating a flexible or minimal relationship with their associated bacterial communities.

Key words: symbiosis, high-throughput sequencing, 16S rRNA metabarcoding, microbiome absolute abundance

Introduction

Microbiomes play a crucial role in the biology of host organisms, influencing everything from individual life histories to population dynamics and broader community interactions (Łukasik & Kolasa, 2024). The diverse microbial communities associated with insects perform key functions, such as supporting nutrition, enhancing immune defense, and facilitating reproduction. These microbial symbionts can shape the evolutionary trajectories of insect hosts and have significant ecological implications across multiple species (McFall-Ngai et al., 2013; Sudakaran et al., 2017). However, not all insects possess substantial microbiomes, and some species seem to thrive with minimal or no reliance on microbial partners (Hammer et al., 2019). Furthermore, even within a single species, microbial diversity and abundance can vary dramatically, indicating that microbiome-host relationships are far more dynamic and context-dependent than previously thought.

Despite a growing pool of projects focusing on insect microbiomes (Thompson et al., 2017), partially due to the simultaneous significant development of sequencing techniques and a decrease in their costs, we are still far from understanding factors shaping the dynamics between hosts and their symbionts. Especially considering the spatio-temporal variation in symbionts' occurrence and quantities in the context of biotic factors such as trophic dependencies, sexes, geographic localization, and genetic background of the hosts. This lack of comprehensive knowledge limits our conclusions on insect biology, including the factors that govern their behavior and ecological roles. Those aspects are crucial in the ongoing mass extinction event and rapidly accelerating climate change (Høye, 2020; Wagner et al., 2021). Fortunately, nowadays, we are equipped with both molecular and bioinformatic toolsets to address these issues more thoroughly and cost-effectively. With proper experimental settings and awareness of caveats related to the applied tools, we can reliably reconstruct insects' microbiomes and conclude about factors shaping them.

Mosquitoes (family Culicidae) have become a focal point for microbiome research, as these insects are not only a nuisance in nearly every region of the world but also critical vectors of many significant diseases (Dada et al., 2021; Guégan et al., 2018; Salgado et al., 2024). Recent studies have shown that mosquito microbiomes may offer novel solutions for controlling hosts' populations and reducing their ability to transmit pathogens ((Capone et al., 2013; Damiani et al., 2010; Hegde et al., 2015)). The potential for manipulating mosquito microbiomes has expanded our knowledge of disease control and mosquito management strategies. Nevertheless, much of this research is focused on a few model species and laboratory populations, leaving the vast majority of the more than 3,000 described mosquito species underexplored with respect to their microbiota (Petersen et al., 2024).

The objective of this study was to comprehensively characterize the microbial communities of mosquitoes across Greenland, a region with relatively limited entomofauna, allowing us to reconstruct ecological networks efficiently. Specifically, we aimed to investigate the mitochondrial genetic diversity of mosquitoes from four distinct regions of Greenland and assess whether Cytochrome Oxidase subunit I (COI) amplicon data could provide additional insights into their evolutionary history and significant

biotic interactions. Additionally, by amplicon-based sequencing of bacterial 16SV4 rRNA gene fragments, we sought to compare the abundance of microbiota in Greenland mosquitoes with those found in other insects from Greenland and in other mosquito species. Our primary focus was to identify differences in microbiota profiles among mosquito populations, habitats, sampling periods, and sexes, while also pinpointing the core microbiome. To accomplish this, we sequenced insect and bacterial marker genes from a collection of 427 specimens obtained from four regions of Greenland.

Methods

Sampling and specimen selection

Insects used in the study (Fig. 1a,b) were collected by sweep netting or Malaise traps (Fig. 1c) in four areas of Greenland: Zackenberg Valley, Constable Pynt, Narsarsuaq, and Nuuk/Kobbefjord, between 2009 and 2023 (Fig. 1d, Table S1). Specimens were preserved in 96% ethanol and stored at -20°C until processing.

We processed large numbers of diverse insects from our Zackenberg 2021 and Kobbefjord 2022 collections, including substantial numbers of mosquitoes (Buczek, Kolasa et al., in prep.), without their prior identification or sexing, and we selected mosquito specimens based on the barcode [see below]. From the remaining multi-species community samples, we specifically selected mosquitoes based on morphology, matching them to the genus *Ochlerotatus* (according to Dahl et al., 2015) and recording the sex of the selected specimens. We selected four additional species from the Zackenberg 2021 collection based on barcode assignation and microbial load estimation. These species were used for microbial abundance comparison along with *Wolbachia*-infected and *Wolbachia*-free *Aedes aegypti* mosquitoes from a long-term laboratory culture at Jagiellonian University and overwintering *Culex pipiens* mosquito females collected in a basement of a cottage in the village of Przybysławice near Miechów, South Poland.

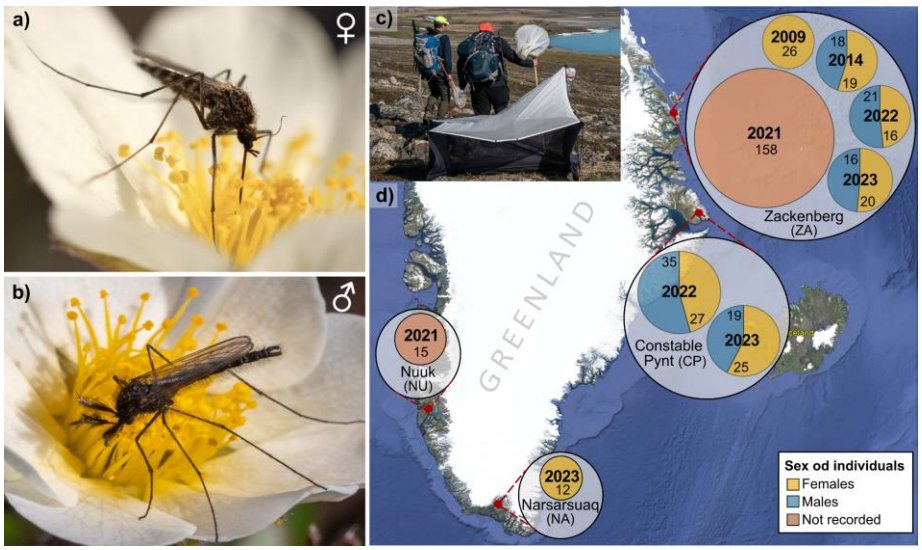

Fig. 1. *Ochlerotatus nigripes* **(a)** female and **(b)** male, feeding on nectar of Dryas flowers in Zackenberg Valley. Photos: PŁ. **(c)** Two of the study authors (TR & MK) in Zackenberg Valley, and the sampling tools: a Malaise trap and a sweep net. **(d)** Mosquito sampling sites and years and sampling depth.

Amplicon library preparation and sequencing

Mosquitoes were processed individually, in different batches (Table S1). The detailed protocols for DNA extraction and library preparation are provided at https://www.protocols.io/view/quantitative-multi-target-amplicon-library-prepara-cwsexebe.

In short, the mosquitoes were placed in 2ml tubes with 200ul' lysis buffer containing proteinase K, homogenized on the Omni Bead Raptor ELITE homogenizer using ceramic beads, and then incubated at 55 °C degrees for two hours. DNA from 40ul homogenate aliquot was purified with SPRI beads. Prior to purification, to the homogenate, we added 20,000 copies of a linearized plasmid carrying an artificial 16S rRNA amplification target, either Ec5502 or Ec5001 (Tourlousse et al., 2016) to enable subsequent microbiome quantification. For each batch of 88 samples, we used a positive extraction control (DNA of a planthopper *Pyrops viridirostris*) and two negative controls (blanks).

A two-step PCR approach was used for amplicon library preparation (Table S2). During the first step, we simultaneously amplified a 418 bp portion of insect cytochrome c oxidase I (COI) gene and 16S rRNA V4 region (16S hereafter) of their associated bacteria, using primers BF3-BR2 and 515F-806R, respectively, with Illumina adapter tails. During the second step, the PCR I products were cleaned with SPRI beads, and a second round of PCR reactions was carried out using a unique combination of indexing primers for each sample. Two positive PCR controls (*Aphrodes markovi*, *Drosophila yakuba*, *Shelfordella lateralis*, or *Pyrops viridirostris* template DNA) and two blank controls were included along with mosquito samples for PCR I, and an additional blank control was included for the second round of PCR reactions. Amplification success and amplicon size were verified by agarose gel electrophoresis after each PCR. Finally, all amplicon libraries were pooled together based on gel band intensity, cleaned, and submitted for sequencing (Table S1).

Initial processing of data

Demultiplexed fastq files for all samples were run through the lab's customized bioinformatic pipeline described in detail at https://github.com/Symbiosis-JU/Bioinformatic-pipelines. Initially, reads were split into bins corresponding to different targeted regions (COI or 16S rRNA) based on primer sequences at the beginning of each read. For each marker gene, paired-end reads were quality-filtered and merged using PEAR (Zhang et al., 2014). Representative sequences were then dereplicated, denoised, taxonomically classified, and clustered into Operational Taxonomic Units (OTUs) based on 97% similarity threshold with VSEARCH (Rognes et al., 2016). Taxonomy assignation was based on the MIDORI reference database version GB239 (Leray et al., 2022) for COI and based on the SILVA database v138 (Quast et al., 2012; Yilmaz et al., 2014) for 16S rRNA sequences.

The resulting OTU tables were then used as input for a series of custom steps required to analyze different target genes and decontamination, as described in detail for each marker gene below.

Mitochondrial diversity analysis

The obtained COI ASV table was used as input for a custom Python script that facilitated barcode recovery, screened for bacterial signal, and provided information on the abundance of dominant genotypes across libraries. These dominant genotypes were designated as a specimen's barcodes if their relative abundance within individual libraries was greater than 60%. Libraries with fewer than 100 total reads were excluded from further analysis.

To analyze ASV diversity within libraries, unassigned ASVs or those identified as chimeras were compared against the NCBI database using the blastN tool. Any ASVs that matched a different insect genus or had less than 90% coverage and identity with mitochondrial sequences of the *Ochlerotatus* genus were discarded. ASVs with fewer than 100 reads in at least one library were categorized as "Others," while dominant ASVs that were represented in only one library were classified as "singletons." The relative abundances of these ASVs were visualized using R Studio (version 2023.06.1+524).

Individual barcodes and metadata on sampling sites and years were used to construct a Minimum Spanning Network in PopArt (version 1.7).

Microbial diversity analyses

The obtained ASV table for the 16S rRNA marker gene was used as a template for a customized decontamination process. We compared a relative abundance of each ASV between experimental samples and blanks used at every laboratory step (extraction, first PCR, second PCR). For each unique 16S rRNA ASV identified through sequencing, we compared the maximum relative abundance it attained across all the negative controls and the experimental libraries. A unique genotype was assigned as a contaminant unless the maximum relative abundance attained in at least one experimental library was more than ten times the maximum relative abundance it attained in any blank library. The remaining sequences were classified as corresponding to 'symbionts'. We exclude any ASV taxonomically assigned as mitochondria, chloroplast, Eukaryota, Archaea or chimera. Additionally, we screened for ASVs classified taxonomically as *Polaromonas* sp., *Brachybcterium* sp. or *Variovorax* sp., which were classified as symbionts despite being known contaminants in the reagent used in our laboratory (Fig. S1, Table S3). Next, 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 (20,000), and the inverse proportion of insect homogenate used during the DNA extraction (5). The reconstructed absolute abundance values were then plotted using ggplot2 (Wickham, 2016), and statistical differences among collection sites, years, and sexes were tested using an ANOVA test in R Studio.

To compare relative abundances and reconstruct the core microbiome, we additionally filtered the dataset to exclude samples with read numbers close to negative controls. Specifically, we set a reads threshold based on the count of symbiont reads present in negative controls after filtering, and experimental samples with lower reads were discarded from further analyses. On average, the remaining 140 negative controls (70 PCR and 70 extraction blanks) had 33 reads. Hence, we decided

that after decontamination with less than 200 reads, every experimental library should be excluded from further analysis. For such a dataset, we grouped 16S rRNA variants based on their higher taxonomic classification (order, genus) to conduct a relative abundance comparison between sites and years of collection using the Phyloseq package (McMurdie & Holmes, 2013), and results were visualized using the ComplexPheatmap package (Gu, 2022; Gu et al., 2016) in R Studio. Additionally, we performed differential microbial abundance analysis using the ANCOMBC package (Lin & Peddada, 2020). Finally, we used the mia (Ernst et al., 2024) and microbiome (Lahti & Shetty, 2012-2019) packages to draw out the core microbiome across samples, and we used ADONIS2 from the vegan package (Oksanen et al., 2024) to compare the composition of the microbial communities among libraries and Principal coordinates analyses (PCoAs) was performed on the weighted unifrac and Bray Curtis distance matrices. We also tested for differences in variance dispersion across groups and performed PERMANOVA to detect statistically significant differences across microbial communities.

Results

COI diversity reveals the geographic distribution patterns of two species of *Ochlerotatus* mosquitoes

We recovered a total of 427 mosquito specimens from our Greenland collection (Fig. 1d) and processed them for multitarget amplicon sequencing. After filtering the data based on the total number of reads, we retained COI data from 383 specimens collected across various years in Zackenberg (n = 266), Constable Pynt (n = 92), Narsarsuaq (n = 10), and Nuuk (n = 15). In total, we retrieved 17,486,481 COI reads, from which 99.6% (17,412,464 reads) were assigned a taxonomy corresponding to the *Ochlerotatus* genus, while the remaining 72,273 reads (0.41%) corresponded to arthropod or mammalian taxa. On average, each mosquito library contained 45,652 COI reads, with a range spanning from 357 to 188,130 reads per sample.

Our analysis identified 2,452 amplicon sequence variants (ASVs), which were classified into 16 operational taxonomic units (OTUs) taxonomically assigned to the genus *Ochlerotatus*. Interestingly, over 75% of the reads within most libraries represented a single COI genotype, which we considered a barcode. One dominant *Ochlerotatus* genotype accounted for 99.72% of the total OTU abundance, with 30 dominant ASVs comprising 96.9% of the reads across the 383 samples. Through comparison with the NCBI database, we identified two species: *Ochlerotatus nigripes* (93.28%) was represented by a barcode in 371 individuals (96.87%), while *Ochlerotatus impiger* (3.63%) was identified in 12 individuals (3.13%) (Fig. 2a).

The most common *O. nigripes* barcode was detected in 318 individuals (83.5%) from various sampling sites and dates, along with multiple low-abundance barcode genotypes that differed from the dominant one by a single nucleotide position out of 418 compared. Additionally, we identified 25 less common *O. nigripes* COI variants. Of these, eight variants constituted more than 75% of the reads in 36 libraries, while 17 variants labeled as ‘singletons’ were present at over 75% abundance in individual

libraries. For *Othe . impiger*, the most abundant variants were represented by five distinct mitochondrial genotypes present across eight libraries, while four additional *O. impiger* 'singleton' variants made up more than 75% of the reads in one library each. These variants differed by up to 7 base pairs, and individuals carrying them were found only at the S/SW sites of Narsarsuaq and Nuuk (Fig. 2a,d). Interestingly, we observed that the primary *O. impiger* barcode (ASV OI.1) was also present at low abundance (ranging from <1% to 12.59%) in 220 *O. nigripes* individuals. This suggests the potential for low-level heteroplasmy within these mosquitoes.

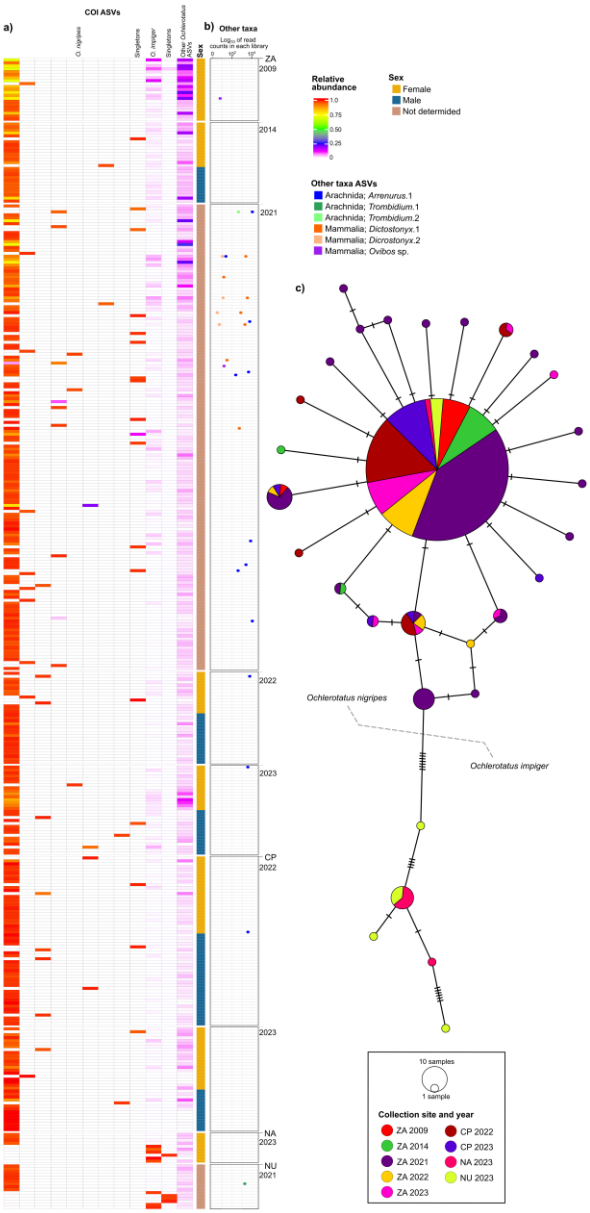

Fig. 2. Mitochondrial diversity of *Ochlerotatus* mosquitoes and biotic interactions. **(a)** Distribution of COI genotypes (ASVs) among *Ochlerotatus* individuals across different sampling years and locations. Columns show the most abundant *Ochlerotatus* ASVs across libraries and the less abundant (<25% per library) in the category "Other ASVs"; additional columns show the specimen's identity based on barcode and sex assignation by color code. Rows represent individual mosquitoes with, at least, 100 reads per library. **(b)** Log transformed COI reads corresponding to different taxa within *Ochlerotatus* libraries; colors represent different ASVs. **(c)** Minimum spanning network (epsilon = 0) of COI haplotypes of *Ochlerotatus* mosquitoes. Each circle represents a unique haplotype, and each hatch mark represents a 1 bp difference between haplotypes. The colors represent the site and year in which individuals were collected, and the size of each circle is proportional to the number of individuals with that haplotype. The dashed line indicates the distribution haplotype division between the two identified species: *O. nigripes* and *O. impiger.*

We also explored the 0.41% of COI reads that did not correspond to *Ochlerotatus* to address possible biotic interactions, such as parasitism or vertebrate bloodmeals. Among these, we identified seven ASVs representing four distinct taxonomic groups (Fig. 1b). The most abundant of these corresponded to a mite from the *Arrenurus* genus, accounting for 60,204 reads (81.3% of the non-mosquito COI signal) across all libraries. Additionally, two individual mosquitoes carried unique ASVs corresponding to mites from the *Trombidium* genus: *Trombidium.1* with 1,983 reads (2.67%) and *Trombidium.2* with 447 reads (0.6%). Finally, we detected vertebrate COI signal corresponding to two ASVs classified within the *Dicrostonyx* genus with a total of 9,618 reads (13% for *Dicrostonyx1* and 0.04% for *Dicrostonyx2*) in 11 libraries. Additionally, two mosquitoes carried ASVs corresponding to the genus *Ovibos*, though these reads were present at low abundance (21 reads, 0.02%).

Microbial abundance and diversity

We obtained a total of 11,144,485 16S rRNA reads across all libraries, which included blank controls (n = 140), *Ochlerotatus* samples (n = 408), as well as reference libraries corresponding to *Aedes aegypti* (n = 71), *Culex pipiens* (n = 18), and *Arrenurus* mites (n = 1), averaging 19,387 reads per library. Notably, the *Ochlerotatus* libraries had a lower average of 14,867 reads per individual, ranging from 0 to 89,984, which was considerably lower compared to *A. aegypti or C. pipiens* libraries, given the number of processed samples (Table S3). We further observed that the average ratio of 16S to COI read numbers per *Ochlerotatus* library was 4.32, which resulted much lower compared to reference mosquitoes, *A. aegypti* (ratio average of 90.2) and similar to that observed in *Culex* mosquitoes (2.27) that were processed simultaneously with the same protocols.

To refine our dataset, we filtered raw libraries based on the blank controls, as described previously and categorized all 16S rRNA ASVs as either ‘symbiont’ or ‘contaminant’ while retaining read counts corresponding to the artificial spike-in target Ec002 for absolute bacterial abundance estimations (Fig. S1a). Ec002, the most abundant ASV across all libraries, comprised 43.87% of the combined dataset (4,889,126 out of 11,144,485 reads) and averaged 62.06% of the reads in each *Ochlerotatus* library (Fig. S1c,d). Among the non-spike-in ASVs, *Brachybacterium* was the most abundant, comprising 17.9% of the initial number of reads per *Ochlerotatus* library (ranging up to 84.38%). However, it also made up significant proportions of the extraction negative controls (15.73%) and the PCR negative controls (62.82%) (Fig. S1a,b, Table 1). We identified *Brachybacterium* as a known reagent contaminant from the Qiagen Multiplex PCR mastermix, and we filtered it from the dataset. Other ASVs corresponding to *Variovorax* sp. were similarly widespread across both negative controls and experimental libraries, with comparable or even greater relative abundances in negative controls. As a result, these ASVs were classified as contaminants according to our criteria (Table S1b). Together, these putative contaminants accounted for an average of 25.63% of non-spike-in reads across the *Ochlerotatus* libraries.

Table 1. Relative abundance of the most abundant ASVs, by category, in raw libraries

	Extraction controls n = 70		PCR controls n = 70		*Ochlerotatus* libraries n = 408	
Category	**Average**	**Maximum**	**Average**	**Maximum**	**Average**	**Maximum**
Spike-in	76.05	100	0.04	100	62.06	95.22
Brachybacterium sp.	15.73	87.09	62.82	92.46	17.90	84.38
Variovorax sp.	2.66	7.43	11.24	55.34	2.76	33.09
Other contaminants	5.32	100	25.38	100	4.98	60.97
Symbionts	0.24	3.04	0.53	1.55	12.30	96.02

Quantitative 16S rRNA amplicon analysis reveals a low microbial load in *Ochlerotatus* mosquitoes

Prior to DNA purification, 20,000 spike-in plasmid copies were added to one-fifth of each insect homogenate sample that would be subsequently sequenced and absolute abundance was estimated based on the spike-in to symbiont read ratio. Libraries with zero spike-in or symbiont reads were discarded from the analysis, and the absolute abundance of 16S rRNA copies was calculated for 399 samples with the symbiont:spike-in ratio averaging 1.19 (range: 0.003 - 153). Assuming no PCR bias, this resulted in an average estimate of 118,734 symbiont 16S rRNA copies per mosquito, with a wide range between 372 and 15,300,000.

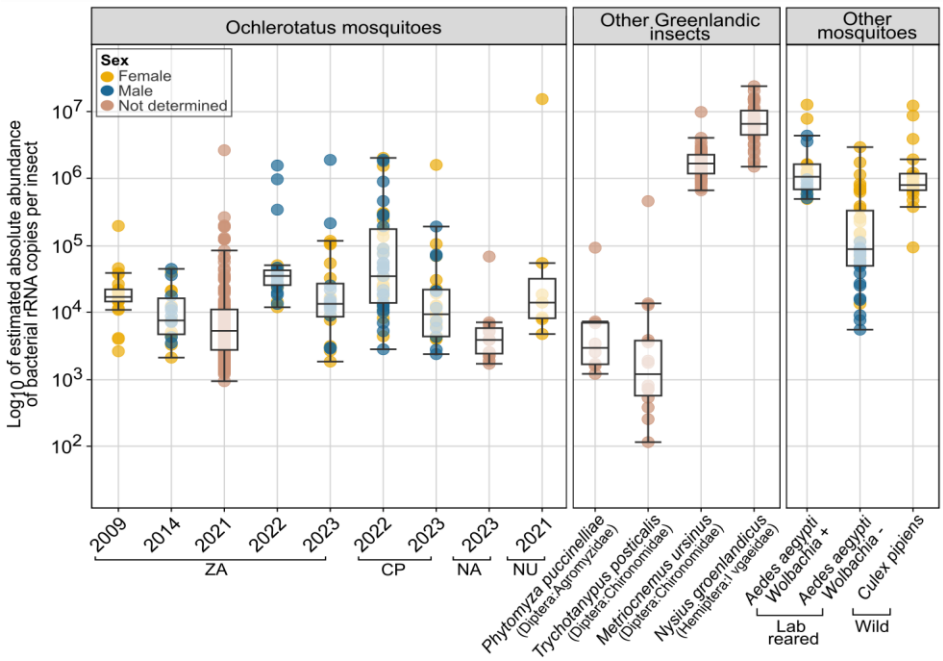

Fig.3. Absolute microbial abundance in Ochlerotatus mosquitoes based on 16S rRNA spike-ins. Boxplots of the median of log-transformed absolute abundance values across mosquito specimens from different sampling years and locations in Greenland, as well as different insects of comparable size sampled in Zackenberg in 2021. As additional references, we quantified microbial abundance on *Wolbachia*-infected and non-infected *Aedes aegypti* reared under controlled conditions in a laboratory, and wild *C. pipens* captured in Poland.

The microbial load of *Ochlerotatus* mosquitoes varied significantly, with microbial abundance ranging from orders of magnitude of 10^2 to 10^7 16S rRNA copies. Notably, 203 mosquitoes (51%) contained fewer than 10^3 copies, while the remaining samples ranged between 10^4 and 10^6 copies (Fig. 3). This microbial load was similar to other insects from Greenland with low microbial abundance but

was one to two orders of magnitude lower than *A. aegypti* mosquitoes, with or without *Wolbachia* endosymbionts, and overwintering female *Culex pipiens.*

An ANOVA test revealed significant variations in microbial load across sampling sites and years ($p < 0.001$), but no significant differences were observed between sexes, either alone or in interaction ($p = 0.587$ and 0.935, respectively). Post hoc Tukey tests highlighted differences between Zackenberg 2014 and Constable Pynt 2022 and between Zackenberg 2014 and 2023 (Table S4).

Relative abundance comparisons reveal a high bacterial diversity with inconsistent distribution patterns across mosquito populations

After excluding contaminants, an average of 1,828 reads per Ochlerotatus library were retained and classified as putative symbionts. However, only 13 of the 408 libraries (3.19%) contained more than 10,000 reads, while 67% of the libraries had fewer than 1,000 reads. Based on these results, a threshold of 200 reads per library was set for relative abundance analyses, ensuring this exceeded the average symbiont abundance detected in control samples (average of 33 symbiont reads). This filtering resulted in the retention of 293 libraries for further analysis.

Microbial community profiles, after filtering for contamination, revealed 15,357 ASVs grouped into 4,162 OTUs, 1,120 genera, and 263 orders. Among these, 11,918 ASVs, 3,222 OTUs, 994 genera, and 242 orders reached at least 1% relative abundance in at least one individual, with the remaining considered very rare due to the overall low absolute abundance of mosquito microbiota. 17 orders were present in at least 10% of the libraries, collectively accounting for 64.14% of the total abundance (Fig. 4b). Despite this, most of the microbial abundance was distributed across 227 other orders, suggesting that while mosquitoes harbor a high diversity of bacteria, most are present in very low abundance, often less than 1% per library. Within these orders, 14 genera representing at least 10% of the libraries comprised 29.6% of total genera abundance (Fig. 4c). *Pseudomonas* was particularly abundant, accounting for 7.7% of total genera abundance, with three *Pseudomonas* ASVs being among the most abundant across all libraries (Fig. 4d). Additionally, 378 *Pseudomonas* ASVs were detected, with the ten most abundant accounting for 55% of the total *Pseudomonas* genus abundance.

Interestingly, libraries dominated by taxa such as *Pseudomonas, Hymenobacter*, and *Bartonella*, which represented more than 75% of the individual library reads, showed higher bacterial absolute abundances. However, upon further inspection, there was no significant correlation between bacterial diversity and absolute 16S rRNA copies within individual libraries (Fig. S2). Ordination analyses using weighted UniFrac and Bray-Curtis metrics did not reveal specific distribution patterns across order, genus, or ASV levels (Fig.5a-c, Fig. S3), although PERMANOVA analyses indicated significant differences among collection sites and years (Table S5). Furthermore, no core microbiome was detected across sampling sites or years above a 0.1% abundance threshold in at least 75% of the libraries (Fig.5d). ANCOM testing confirmed that while some taxa were shared across sites and years, this applied to fewer than 50% of libraries, and these shared taxa exhibited no clear distribution patterns (Table S6).

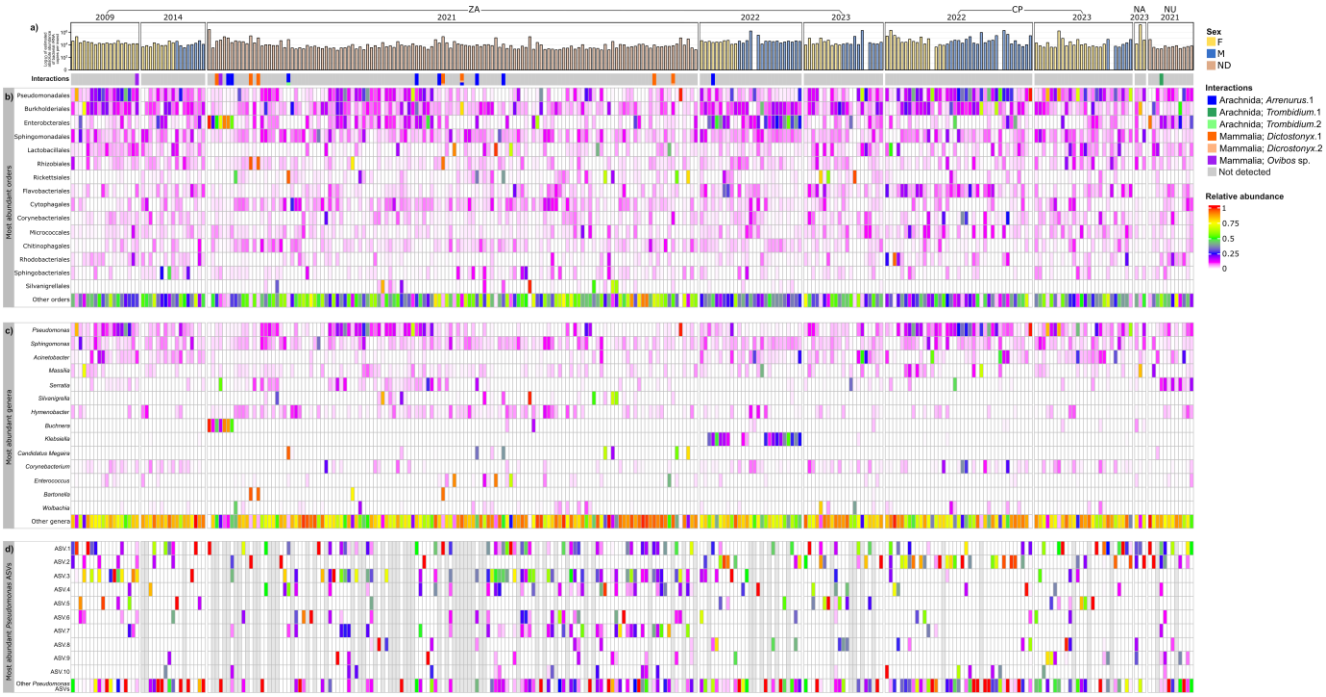

Fig. 4. Bacterial diversity across Ochlerotatus individuals. **(a)** Absolute microbial abundance was quantified in individual mosquitoes across locations and years; bar colors indicate the sex of specimens. Biotic interactions (i.e. parasitism, bloodmeals) are shown by color code associated taxa and bacterial relative abundance is shown for the most abundant taxa at the level of **(b)** order, **(c)** genus, and **(d)** ASV across libraries.

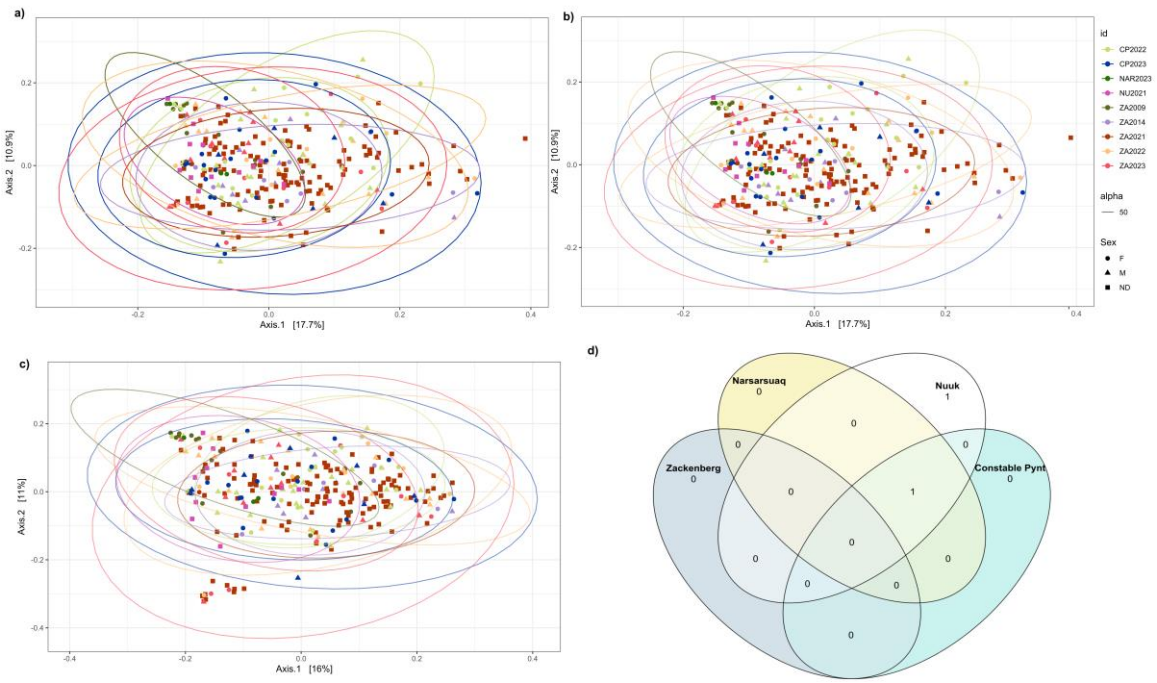

Fig. 5. Microbial community composition among sampling sites and mosquito sexes. PCoAs showing weighted Unifrac distances at the level of **(a)** order, **(b)** genera, and **(c)** ASVs. The colors represent different sampling years and sites; the shapes represent the sex of the specimens. **(d)** Core microbiome across sampling sites at a bacterial detection level of 0.1 % in at least 75 % of the analyzed libraries.

Discussion

The last few decades helped us to understand better the role of microbes in the evolution and adaptation to various environments and habitats of insect hosts (Łukasik & Kolasa, 2024). Although we have begun

to comprehend the intricate relation between insect hosts and their microbes due to the development of bacterial metabarcoding, the knowledge about factors driving symbiotic dynamics in natural insect communities is virtually non-existent. Especially considering that the most widely used (bacterial marker gene amplicon sequencing) is burdened with technical caveats (Knight et al., 2018) that usually are not adequately addressed and might blur the conclusions about microbiome dynamics. Hence, it is necessary to carefully design experiments on the one hand and use proper tool sets for addressing particular hypotheses on the other if the scientific community aims to understand and mitigate the mechanisms behind global insect decline (Høye, 2020; Wagner et al., 2021). For a long time, insects have been thought to require more or less stable microbiomes to function or develop properly (Gupta & Nair, 2020). However, we have growing evidence that some insect species facing intense environmental pressure have relatively simple and low-abundant microbiomes (Buczek, Kolasa et al., 2024), which leaves the question of how to study virtually sterile insect species and what makes them immune to omnipresent bacteria colonization.

Here, by meticulously analyzing insect COI and bacterial 16S marker genes, we offer new insights into the geographic distribution patterns and microbial diversity of two *Ochlerotatus* mosquito species from Greenland. Our insect marker gene analyses led us to identify two species across the studied area: *Ochlerotatus nigripes* and *Ochlerotatus impiger*. The first one was the predominant species, accounting for a vast majority (96.87%) of individuals across various sites and years. At the same time, *O. impiger* was less abundant and confined to the southern and southwestern regions (Narsarsuaq and Nuuk). Such marked geographic segregation of the two species might reflect differences in environmental or ecological preferences emerging as barriers (Renault et al., 2018) or might be a sign of ongoing range expansion of *O. impiger* due to climate change (Hill et al., 2011). Distribution monitoring spread further into the future with a simultaneous comparison of historical samples should provide an answer to this question.

Intriguingly, low-level heteroplasmy was detected within *O. nigripes* individuals, evidenced by the occasional presence of *O. impiger* barcodes at low abundances. Although these findings contrast with a recent study reporting the absence of heteroplasmy in the genome of a mosquito species, this discovery raises questions about the potential for interspecies genetic exchange or incomplete lineage sorting within the genus. Alternatively, this faint signal might result from environmental contamination, as airborne mosquito DNA is ubiquitous in Greenlandic samples (Kolasa et al., in prep). However, these COI variants found in *O. nigripes* individuals from all sampled sites corresponded to representative barcodes from *O. impiger* species, which were exclusively found in the south of Greenland. Since they were stored and processed separately and sequenced on different lanes, we can confidently rule out cross-contamination between samples. Nonetheless, further investigation into the mechanisms behind this heteroplasmy could reveal important evolutionary dynamics between these closely related species.

Our microbial analysis revealed significant findings regarding the bacterial load and diversity in Greenlandic mosquitoes. Consistent with other Greenland insects of comparable size (Buczek/Kolasa et al., in prep), *Ochlerotatus* mosquitoes exhibited a notably low microbial load, averaging 118,734

symbiont 16S rRNA copies per mosquito, which is one to two orders of magnitude lower than in tropical species like *Aedes aegypti* or overwintering *Culex pipiens* collected in Poland. Although mosquito microbiota origin and acquisition have been debated (Guégan et al., 2018; Minard et al., 2013), evidence suggests its environmental origin during the larval life stage (Coon et al., 2017; Dada et al., 2021). The differences in bacterial abundance between Greenlandic, tropical, and Polish mosquitos might reflect differences in their guts as a suitable habitat for bacterial colonization. Conversely, Greenlandic mosquitos might lack the nutrients required to sustain a significant load of bacteria in their guts. Nutrition availability or the availability of environmental bacteria capable of colonizing mosquitos might be supported by our results demonstrating considerable variability in microbial load across different sampling sites and years, as indicated by significant differences in microbial abundance ($p < 0.001$).

On the other hand, although some libraries showed dominance by specific taxa, such as *Pseudomonas*, *Hymenobacter*, and *Bartonella*, there was no significant correlation between microbial diversity and total bacterial load across individual mosquitoes. This suggests that random processes may influence bacterial colonization rather than specific factors related to host microbial capacity. Furthermore, ordination analyses failed to reveal consistent distribution patterns of bacterial communities across mosquito populations, and no core microbiome was identified above a 0.1% abundance threshold, which opposes the finding in different mosquito species (Dada et al., 2021; Hegde et al., 2015) These claims contradict our ANCOM testing results, which showed that while some bacterial taxa are shared across sampling sites and years, fewer than 50% of libraries contained these shared taxa, and their distribution exhibited no clear patterns. Our results add to the growing evidence that not all animals need microbiomes (Hammer et al., 2019), and microbes colonizing mosquitos can be described as ‘transient’ and might be irrelevant to host functioning. However, the question of what are mechanisms underlying insect sterility and how common it is across insect taxa requires further investigation.

In conclusion, our study highlights the complex interactions between geography, host species, and microbial communities in Greenlandic mosquitoes. The dominance of *O. nigripes* across most regions contrasts with the restricted distribution of *O. impiger*, and while mosquitoes host a diverse bacterial community, the overall microbial load remains low and varies significantly across sites and years. These findings suggest that the extreme arctic environment may play a crucial role in shaping the microbiota of these mosquitoes. Future research should explore the potential adaptive significance of these microbial communities.

Acknowledgements

We thank Ewa Chrostek and the members of the Chrostek lab for providing the *A. aegypti* specimens and for generously offering their facilities, resources, and expertise, including the laboratory space, equipment, and supervision necessary for the rearing and maintenance of the specimens.

This project was supported by the Polish National Agency for Academic Exchange grant no. PPN/PPO/2018/1/00015, and Polish National Science Centre grants no. 2018/31/B/NZ8/01158 and 2021/43/B/NZ8/03376 and has received funding from INTERACT III Transnational Access under the European Union H2020 Grant Agreement No. 629 and 729.
The research has also been supported by a grant from the Priority Research Area BioS under the Strategic Programme Excellence Initiative at Jagiellonian University, we also acknowledge support from the National Genomics Infrastructure in Stockholm for assistance with massively parallel sequencing.

References

Capone, A., Ricci, I., Damiani, C., Mosca, M., Rossi, P., Scuppa, P., Crotti, E., Epis, S., Angeletti, M., Valzano, M., Sacchi, L., Bandi, C., Daffonchio, D., Mandrioli, M., & Favia, G. (2013). Interactions between Asaia, Plasmodium and Anopheles: New insights into mosquito symbiosis and implications in Malaria Symbiotic Control. *Parasites & Vectors*, *6*(1), 182. https://doi.org/10.1186/1756-3305-6-182

Coon, K. L., Valzania, L., McKinney, D. A., Vogel, K. J., Brown, M. R., & Strand, M. R. (2017). Bacteria-mediated hypoxia functions as a signal for mosquito development. *Proceedings of the National Academy of Sciences*, *114*(27). https://doi.org/10.1073/pnas.1702983114

Dada, N., Jupatanakul, N., Minard, G., Short, S. M., Akorli, J., & Villegas, L. M. (2021). Considerations for mosquito microbiome research from the Mosquito Microbiome Consortium. *Microbiome*, *9*(1), 36. https://doi.org/10.1186/s40168-020-00987-7

Damiani, C., Ricci, I., Crotti, E., Rossi, P., Rizzi, A., Scuppa, P., Capone, A., Ulissi, U., Epis, S., Genchi, M., Sagnon, N., Faye, I., Kang, A., Chouaia, B., Whitehorn, C., Moussa, G. W., Mandrioli, M., Esposito, F., Sacchi, L., … Favia, G. (2010). Mosquito-Bacteria Symbiosis: The Case of Anopheles gambiae and Asaia. *Microbial Ecology*, *60*(3), 644–654. https://doi.org/10.1007/s00248-010-9704-8

Ernst F, Shetty S, Borman T, Lahti L (2024). *mia: Microbiome analysis*. R package version 1.13.36, https://github.com/microbiome/mia.

Gu, Z. (2022). Complex heatmap visualization. *iMeta*, *1*(3), e43. https://doi.org/10.1002/imt2.43

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*(18), 2847–2849.

https://doi.org/10.1093/bioinformatics/btw313

Guégan, M., Zouache, K., Démichel, C., Minard, G., Tran Van, V., Potier, P., Mavingui, P., & Valiente Moro, C. (2018). The mosquito holobiont: Fresh insight into mosquito-microbiota interactions. *Microbiome*, *6*(1), 49. https://doi.org/10.1186/s40168-018-0435-2

Gupta, A., & Nair, S. (2020). Dynamics of Insect–Microbiome Interaction Influence Host and Microbial Symbiont. *Frontiers in Microbiology*, *11*, 1357. https://doi.org/10.3389/fmicb.2020.01357

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

Hegde, S., Rasgon, J. L., & Hughes, G. L. (2015). The microbiome modulates arbovirus transmission in mosquitoes. *Current Opinion in Virology*, *15*, 97–102. https://doi.org/10.1016/j.coviro.2015.08.011

Hill, J. K., Griffiths, H. M., & Thomas, C. D. (2011). Climate Change and Evolutionary Adaptations at Species' Range Margins. *Annual Review of Entomology*, *56*(1), 143–159. https://doi.org/10.1146/annurev-ento-120709-144746

Høye, T. T. (2020). Arthropods and climate change – arctic challenges and opportunities. *Current Opinion in Insect Science*, *41*, 40–45. https://doi.org/10.1016/j.cois.2020.06.002

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

Lahti L, Shetty S (2012-2019). Microbiome R package. Bioconductor. https://doi.org/10.18129/B9.bioc.microbiome

Leray, M., Knowlton, N., & Machida, R. J. (2022). MIDORI2: A collection of quality controlled, preformatted, and regularly updated reference databases for taxonomic assignment of eukaryotic mitochondrial sequences. *Environmental DNA*, *4*(4), 894–907. https://doi.org/10.1002/edn3.303

Lin, H., & Peddada, S. D. (2020). Analysis of compositions of microbiomes with bias correction. *Nature Communications*, *11*(1), 3514. https://doi.org/10.1038/s41467-020-17041-7

Łukasik, P., & Kolasa, M. R. (2024). With a little help from my friends: The roles of microbial symbionts in insect populations and communities. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *379*(1904), 20230122. https://doi.org/10.1098/rstb.2023.0122

McFall-Ngai, M., Hadfield, M. G., Bosch, T. C. G., Carey, H. V., Domazet-Lošo, T., Douglas, A. E., Dubilier, N., Eberl, G., Fukami, T., Gilbert, S. F., Hentschel, U., King, N., Kjelleberg, S., Knoll, A. H., Kremer, N., Mazmanian, S. K., Metcalf, J. L., Nealson, K., Pierce, N. E., … Wernegreen, J. J. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences*, *110*(9), 3229–3236. https://doi.org/10.1073/pnas.1218525110

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Minard, G., Mavingui, P., & Moro, C. V. (2013). Diversity and function of bacterial microbiota in the mosquito holobiont. *Parasites & Vectors*, *6*(1), 146. https://doi.org/10.1186/1756-3305-6-146

Oksanen J, Simpson G, Blanchet F, Kindt R, Legendre P, Minchin P, O'Hara R, Solymos P, Stevens M, Szoecs E, Wagner H, Barbour M, Bedward M, Bolker B, Borcard D, Carvalho G, Chirico M, De Caceres M, Durand S, Evangelista H, FitzJohn R, Friendly M, Furneaux B, Hannigan G, Hill M, Lahti L, McGlinn D, Ouellette M, Ribeiro Cunha E, Smith T, Stier A, Ter Braak C, Weedon J (2024). *vegan: Community Ecology Package*. R package version 2.7-0, https://github.com/vegandevs/vegan,

Petersen, V., Santana, M., Karina-Costa, M., Nachbar, J. J., Martin-Martin, I., Adelman, Z. N., & Burini, B. C. (2024). Aedes (Ochlerotatus) scapularis, Aedes japonicus japonicus, and Aedes (Fredwardsius) vittatus (Diptera: Culicidae): Three Neglected Mosquitoes with Potential Global Health Risks. *Insects*, *15*(8), 600. https://doi.org/10.3390/insects15080600

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

Renault, D., Laparie, M., McCauley, S. J., & Bonte, D. (2018). Environmental Adaptations, Ecological Filtering, and Dispersal Central to Insect Invasions. *Annual Review of Entomology*, *63*(1), 345–

368. https://doi.org/10.1146/annurev-ento-020117-043315

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

Salgado, J. F. M., Premkrishnan, B. N. V., Oliveira, E. L., Vettath, V. K., Goh, F. G., Hou, X., Drautz-Moses, D. I., Cai, Y., Schuster, S. C., & Junqueira, A. C. M. (2024). The dynamics of the midgut microbiome in *Aedes aegypti* during digestion reveal putative symbionts. *PNAS Nexus*, *3*(8), pgae317. https://doi.org/10.1093/pnasnexus/pgae317

Sudakaran, S., Kost, C., & Kaltenpoth, M. (2017). Symbiont Acquisition and Replacement as a Source of Ecological Innovation. *Trends in Microbiology*, *25*(5), 375–390. https://doi.org/10.1016/j.tim.2017.02.014

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J. T., Mirarab, S., Zech Xu, Z., Jiang, L., … Zhao, H. (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature*, *551*(7681), 457–463. https://doi.org/10.1038/nature24621

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

Wagner, D. L., Grames, E. M., Forister, M. L., Berenbaum, M. R., & Stopak, D. (2021). Insect decline in the Anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences*, *118*(2), e2023989118. https://doi.org/10.1073/pnas.2023989118

Wickham, H. (with Sievert, C.). (2016). *ggplot2: Elegant graphics for data analysis* (Second edition). Springer.

Yilmaz, P., Parfrey, L. W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., Schweer, T., Peplies, J., Ludwig, W., & Glöckner, F. O. (2014). The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research*, *42*(D1), D643–D648. https://doi.org/10.1093/nar/gkt1209

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