PNAS Nexus, 2024, 3, pgae317
https://doi.org/10.1093/pnasnexus/pgae317
Advance access publication 1 August 2024
Research Report

The dynamics of the midgut microbiome in *Aedes aegypti* during digestion reveal putative symbionts

João Felipe M. Salgado [a,b,*], Balakrishnan N. V. Premkrishnan [c], Elaine L. Oliveira [c], Vineeth Kodengil Vettath [c], Feng Guang Goh [d,e], Xinjun Hou [d,e], Daniela I. Drautz-Moses [c], Yu Cai [d,e], Stephan C. Schuster [c] and Ana Carolina M. Junqueira [b,*]

[a]RG Insect Microbiology and Symbiosis, Max Planck Institute for Terrestrial Microbiology, Karl-von-Frisch str. 10, Marburg 35043, Germany
[b]Departamento de Genética, Instituto de Biologia, Universidade Federal do Rio de Janeiro, 373 Avenida Carlos Chagas Filho, Rio de Janeiro, RJ 21941-902, Brazil
[c]Singapore Center for Environmental Life Sciences Engineering (SCELSE), Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore
[d]Temasek Life Sciences Laboratory, National University of Singapore, 1 Research Link, Singapore 117604, Singapore
[e]Department of Biological Sciences, National University of Singapore, 16 Science Drive 4, Block S3, #05-01, Singapore 117558, Singapore
*To whom correspondence should be addressed: Ana Carolina M. Junqueira, Email: anacmj@gmail.com; João Felipe M. Salgado, joao.salgado@mpi-marburg.mpg.de
Edited By Christopher Dupont

Abstract

Blood-feeding is crucial for the reproductive cycle of the mosquito *Aedes aegypti*, as well as for the transmission of arboviruses to hosts. It is postulated that blood meals may influence the mosquito microbiome but shifts in microbial diversity and function during digestion remain elusive. We used whole-genome shotgun metagenomics to monitor the midgut microbiome in 60 individual females of *A. aegypti* throughout digestion, after 12, 24, and 48 h following blood or sugar meals. Additionally, ten individual larvae were sequenced, showing microbiomes dominated by *Microbacterium* sp. The high metagenomic coverage allowed for microbial assignments at the species taxonomic level, also providing functional profiling. Females in the post-digestive period and larvae displayed low microbiome diversities. A striking proliferation of *Enterobacterales* was observed during digestion in blood-fed mosquitoes. The compositional shift was concomitant with enrichment in genes associated with carbohydrate and protein metabolism, as well as virulence factors for antimicrobial resistance and scavenging. The bacterium *Elizabethkingia anophelis* (Flavobacteriales), a known human pathogen, was the dominant species at the end of blood digestion. Phylogenomics suggests that its association with hematophagous mosquitoes occurred several times. We consider evidence of mutually beneficial host-microbe interactions raised from this association, potentially pivotal for the mosquito's resistance to arbovirus infection. After digestion, the observed shifts in blood-fed females' midguts shifted to a sugar-fed-like microbial profile. This study provides insights into how the microbiome of *A. aegypti* is modulated to fulfil digestive roles following blood meals, emphasizing proliferation of potential symbionts in response to the dynamic midgut environment.

Keywords: whole-genome sequencing metagenomics, digestion, hematophagy, *Enterobacterales*, *Elizabethkingia anophelis*

Significance Statement

This study describes compositional and functional dynamics during digestion of the midgut microbiome of the mosquito vector *Aedes aegypti*. A comprehensive metagenomic dataset was generated to analyze the midgut microbiome of females after blood or sugar meals. Blood meals stimulate *Enterobacterales* proliferation during its digestion, also inducing a functional shift and suggesting a role in assisting blood metabolism. The metagenomic approach reveals that, in later stages of blood digestion, the human-pathogenic bacterium *Elizabethkingia anophelis* is predominant in the midgut. This association may have implications in the evolution of hematophagy, human pathogenesis, and viral transmission. These findings shed light on potential strategies for insect-borne disease control by harnessing resident microbial components that influence *A. aegypti*'s vectorial capacity and pathogen transmission.

Introduction

The knowledge accumulated in the last decade about the association of multicellular eukaryotes with microorganisms has provided a paradigm shift in what was previously known of metabolic, physiological, and homeostatic fitness in virtually every multicellular host organism (1–3). Microbiome research has primarily focused on humans, but recent advances have expanded the investigation to other animals. In this context, several insect vectors have been targeted by metagenomic approaches, unveiling complex microbiome interactions that regulate processes essential to the host life cycle (4–8). While diverse environmental factors may contribute to variations in the microbiome

Competing interest: The authors declare no competing financial interests.
Received: November 7, 2023. **Accepted:** July 2, 2024

of insects (6) the nutritional source appears to be especially impactful in modulating microbial communities (9, 10). Likewise, the patterns and dynamics of microbial proliferation in holometabolous insects can also be determined by their habitats and developmental stages (11).

In hematophagous mosquitoes that rely on the blood meal for egg development, oviposition, and lifespan (12, 13), the food type and source proved to be fundamental for the interaction with their intestinal microbiota, influencing the vector's susceptibility to viruses and their capacity to transmit pathogens to hosts (14–17). These findings are essential for developing successful strategies for vector control based on microbiota manipulation, such as those reported for *Wolbachia* infections (18, 19). The yellow fever mosquito, *Aedes aegypti*, is the primary vector of viral diseases worldwide, such as yellow fever, Zika, chikungunya, and dengue. The latter, alone, is responsible for a global economic burden of US$9 billion per year (20, 21). Previous studies based on 16S sequencing have reported a core microbiome composed of aerobic and facultative-anaerobic bacteria in *Aedes* spp (22). They show that multiple factors can impact *A. aegypti* microbiota, such as habitat, environmental contamination with fertilizers or antibiotics, sex, developmental stage, or nutrition. Additionally, substantial differences in the bacterial composition were found in the midgut of *A. aegypti* fed on distinct food sources and in mosquitoes fed with blood from different animal hosts (23, 24). Although metagenome-assembled genomes (MAGs) have confirmed the presence of major mosquito-associated symbionts (25), the microbial shifts triggered by the blood meal in the midgut and their functional implications throughout the digestion have not been fully explored by large-scale metagenomic approaches.

In the present study, we provide an in-depth analysis of 70 individual shotgun metagenomes from *A. aegypti*, including 60 adult females fed with either blood or sugar, and ten larvae. The dynamics of the microbiome were monitored in both dietary groups of adults at 12, 24, and 48 hours after feeding ($n = 10$ for each). Our metagenomic approach allowed for the microbial taxonomic assignment up to the species level, revealing a concerted proliferation of enterobacteria in the mosquito's midgut during the blood meal digestion. This compositional shift was accompanied by a functional enrichment in microbial metabolic functions involved in the catabolism of amino acids, sugar, as well as virulence factors related to antimicrobial resistance and scavenging. The sugar-fed group presents a higher diversity in its microbiome when compared to the blood-fed group, but no significant functional enrichment. The post-digestion period in both diet groups is characterized by the presence of the flavobacterium *Elizabethkingia anophelis*. This study documents the occurrence of this symbiotic bacterium in *A. aegypti*, its potential human-pathogenic evolutionary origins and discusses the implications of this association for transmission of arboviruses. We also demonstrate that its presence in the midgut is modulated by diet. The microbiome of different development stages was also evaluated by comparing larval and adult microbiomes, revealing a notable shift in microbiome profiles from the larval to adult stages. Larvae exhibited low microbial diversity, dominated by the potentially transstadial genus *Microbacterium*.

Results and discussion

Metagenomic datasets

The total DNA extraction of the midgut of 60 females of *A. aegypti* provided an average of 1.078 ng/μL ± 0.663, while ten individual larvae in the fourth-instar yielded 0.599 ng/μL ± 0.121. Adult mosquitoes fed with blood showed a lower DNA yield when compared to the mosquitoes fed with sugar, with an average of 0.783 ng/μL ± 0.385 for blood-fed adults and 1.372 ng/μL ± 0.752 for sugar-fed adults. The total DNA amount recovered from each individual sample is shown in Figure S1 and Table S1. Despite the low biomass, the average number of reads generated per-sample was approximately 45 million (negative and environmental controls excluded, Table S1). A total of 3,213,701,600 reads were generated, from which 1,376,183,852 reads (~42.82%) were classified as non-host reads after *in-silico* removal of mosquito genomic sequences. The total DNA and metagenomic reads per-group before and after mapping and the reads generated for controls are in Table S1. A total of 1,164,457,880 paired-end reads were used in the subsequent metagenomic analyses (an average of: fourth-instar larvae = 38,054,610 ± 6,105,856; blood-fed adults = 20,139,425 ± 11,488,999; sugar-fed adults = 5,990,967 ± 624,586).

To assess whether the number of microbial phylotypes in each sample was a function of the sequencing depth, we performed interpolation and extrapolation of the reads using microbial richness at the genus and species taxonomic levels. Fourth-instar larvae and blood-fed adults reach a *plateau* of rarefaction at approximately 400,000 reads. In comparison, sugar-fed adults reached the regression threshold at approximately 50,000 reads, showing that all curves were rarefied to the point where the size of the datasets no longer contributed significantly to the increase of microbial diversity (Figure S2). Therefore, the metagenomic dataset sizes used in this work did not create a bias in the diversity of the microbiomes.

Larval microbiome and experimental controls

Larvae are exclusively colonized by Actinomycetota of the genus *Microbacterium* (3,403,719 reads; 65 phylotypes detected; Fig. 1). They display a significantly higher diversity with the Chao1 index (Fig. 2A) than the adult groups. Still, this result is due to a decreased evenness, which is reflected in the diversity analysis with Simpson reciprocal indices (Fig. 2A). Recently, the role of *Microbacterium* sp. was assessed in axenic *A. aegypti* larvae, showing that it is the only taxon that was not associated with an increased rate of survival to adulthood (26). This finding suggests that *Microbacterium* sp. does not contribute significantly to the development of larvae despite their dominance. Nevertheless, its colonization in *A. aegypti* larvae may be relevant in other aspects, including symbioses that antagonize the establishment of fungal entomopathogens, such as *Metarhizium robertsii* (27). This perspective may explain why fungal pathogens such as *Metarhizium majus* and *Aspergillus flavus* coincide with low *Microbacterium* sp. colonization in adult mosquitoes (Fig. 1B). Additionally, the water sample where larvae were reared did not show the presence of *Microbacterium* sp. (Figure S3), further indicating that its richness in larval samples is not acquired from the environment and is indeed typical of this developmental stage. The actinomycetes *Leifsonia aquatica* (111,018 reads), *Leucobacter chironomi* (160,140 reads), and *Corynebacterium sp.* (223,942 reads) were also found in larvae but not in adult mosquitoes. Transstadial transmission of the bacteria *Pseudomonas* sp. and *E. anophelis* has been previously reported in *Anopheles* spp. and *Aedes triseratus* (28, 29). In this study, we observed a distinct microbiome profile in the larval samples, compared to adult mosquitoes. Blood and sugar-fed adults showed little to no abundances of *Microbacterium*, suggesting that the mosquito *A. aegypti* may carry *Microbacterium* through its life cycle, although significant differences exist in the dominant taxa from larvae to adult mosquitoes.

In addition to using rearing water as an environmental control for the larval microbiome, we also sequenced the blood and sugar

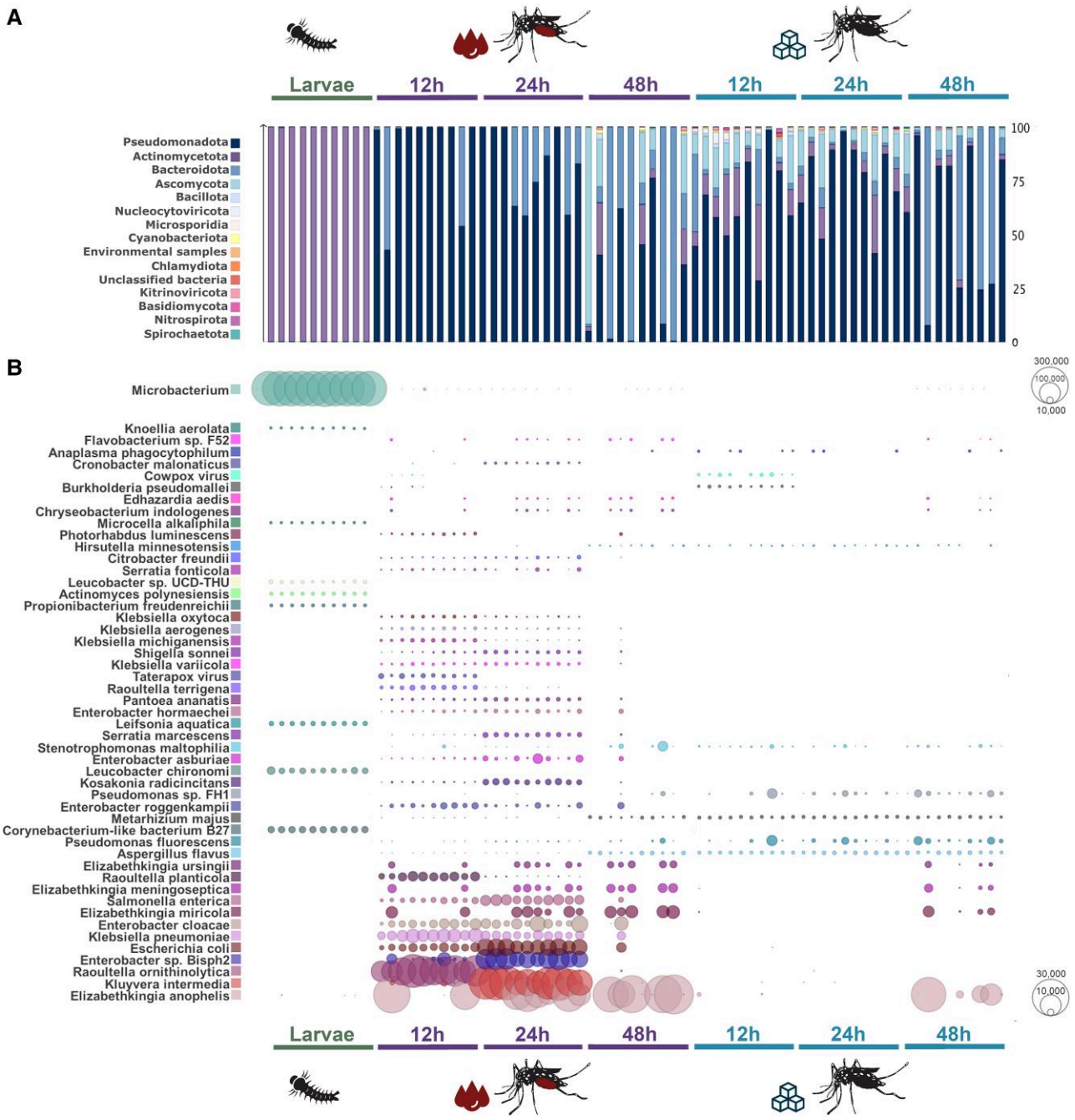

Fig. 1. The microbiome composition of female adults and larval stages. A) Compositional difference at the phylum level. The bins in the bar chart are scaled as percentages. B) Species-level variation of the top 50 phylotypes across adults and larvae. Bubbles display the relative abundance of microbial species on a quadratic scale based on the normalized number of assigned reads for each adult midgut or larval sample. Different species of *Microbacterium* sp. were collapsed to the taxonomic level of the genus and displayed in a separate scale (top-left) to enable the comparison between different taxa in the experimental groups.

solutions used to feed the adults to assess whether these meals were the source of microbial colonization in diet groups. In general, the number of reads attributed to microbial phylotypes was drastically lower in all control samples, corresponding to a decline of 98.6% in assigned reads (268,671 reads on average) compared to the mosquito samples (20,409,956 reads on average). In the blood sample, more than 95% of reads were assigned to *Sus scrofa*, coinciding with the source of the blood used for the mosquitoes' meals. Viral reads were detected in the blood source and were assigned to the Taterapox virus, which was also present in a small portion of the microbiome of adult mosquitoes fed with blood (Fig. 1B and Figure S4). This is indicative that females likely acquired these viruses during the blood meal, albeit they most likely do not infect the mosquito. In the water-sugar solution used for feeding, the bacterial species *Microbacterium* sp. and *Pseudomonas fluorescens* were found with 666 and 441 reads, respectively. In the negative controls (blanks), *Methylobacterium* sp. was detected (34,651 reads; Figure S3). This taxon was previously described as a contaminant in commercial kits commonly used for metagenomics DNA extraction (30). Together, the metagenomic analyses of experimental controls show that our analyses were not influenced by the microbial components previously present in the

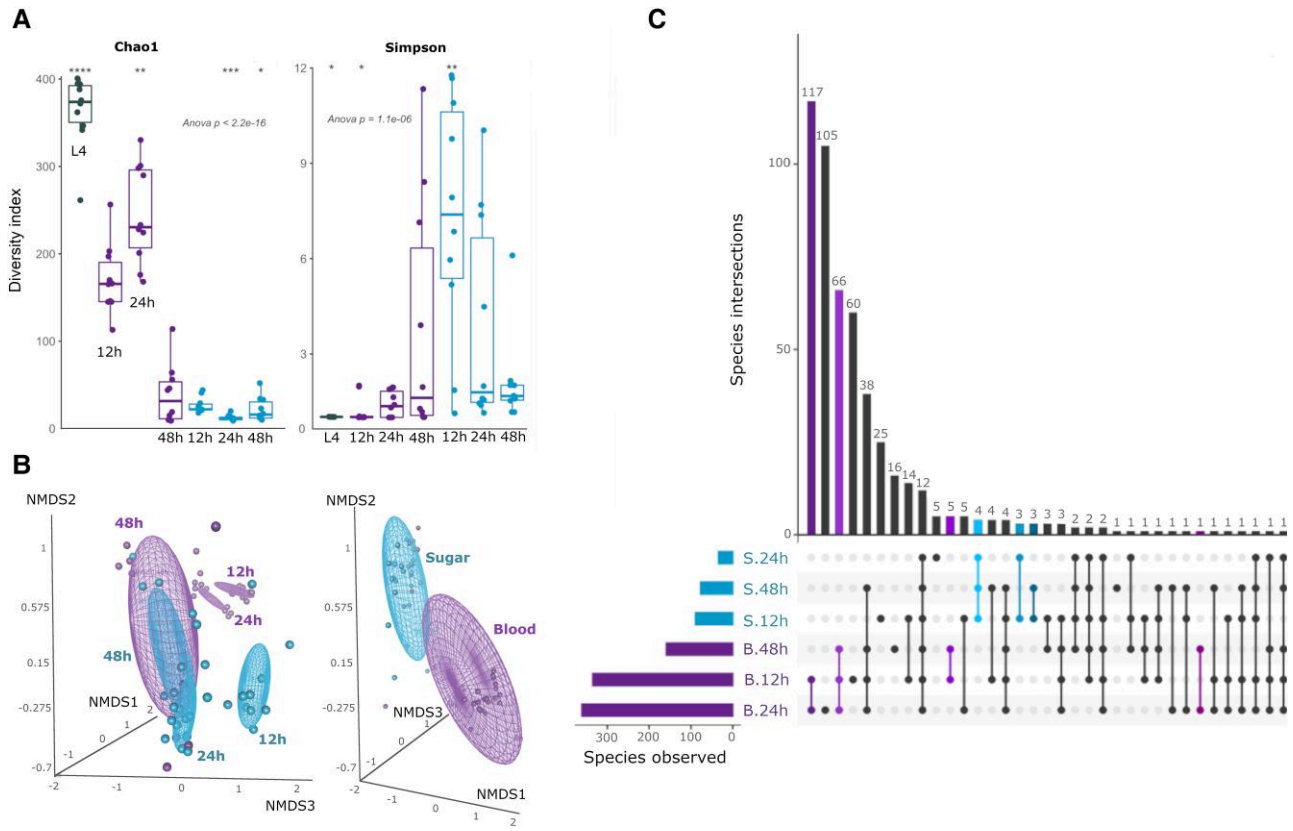

Fig. 2. Diversity profiles of mosquitoes' microbiome in different developmental stages and throughout 48 h of the digestion of different diets in adults. A) Boxplots showing the distribution of diversity indices and entropies. The global and pairwise significances are assessed with ANOVA and Wilcoxon's tests, respectively. B) Tridimensional scatterplot of NMDS using Bray–Curtis dissimilarities (stress = 0.07, model fit = 99%) of sugar and blood-fed groups collected 12, 24, and 48 h after feeding. Ellipses represent confidence intervals of the distances calculated from each group's centroids. ADONIS and ANOSIM tests support the contribution of each dependent variable (type of diet and time elapsed after feeding) in the distribution of microbial species. The different sizes of spheres are a function of the tridimensional perspective. C) Shared and unique microbial species between blood or sugar-fed adult mosquitoes. The intersections (connected dots) are coded by type of diet (B.12 h = Blood 12 h post-feeding; B.24 h = Blood 24 h post-feeding; B.48 h = Blood 48 h post-feeding; S.12 h = Sugar 12 h post-feeding; S.24 h = Sugar 24 h post-feeding; S.48 h = Sugar 48 h post-feeding), as are their correspondent bins in the histogram. The intersections sizes, representing the respective number of shared species between given groups, are displayed in the histogram. Bins and dots of Species shared between different diets are black.

environment where larvae were reared, meal solutions, or in the DNA extraction kit.

Shifts in microbial composition are driven by diet

The composition of the normalized datasets shows that Bacteria is the prevalent microbial domain in larvae and in the midgut of females of *A. aegypti* (approximately 99.6% Bacteria, 0.3% Fungi, 0.05% Viruses, 0.007% Archaea, Figure S4), regardless of the food source. Individually, samples in the sugar-fed group displayed the highest volume of reads assigned to Fungi, irrespective of the digestive period (Figure S4). A per-sample 10 to 15% of fungal reads in sugar-fed adults was composed of the phylum Ascomycota (Fig. 1A). Nevertheless, the bacterial phylum Pseudomonadota composed a large portion of most metagenomes analyzed, with relative abundances ranging from 45 to 95% of the microbiomes in individual samples (Fig. 1A).

At the species taxonomic level, the top 50 microorganisms assigned to each sample are detailed in Fig. 1B. A total of 802 microbial phylotypes were detected in the metagenomes with per-group averages as follows: fourth-instar larvae = 301.5 ± 17.8; blood-fed adults = 166.5 ± 93; sugar-fed adults = 18.5 ± 11. A shift in the microbiome composition was observed in the earliest hours of digestion (Fig. 1). At 12 and 24 h, the group of blood-fed adults displayed a dramatic increase in abundance of the phylum Pseudomonadota, with the proliferation of *Enterobacterales*. At 48 h post blood meal, the microbiome of the blood-fed adults group exhibited yet another shift and *Enterobacterales* was no longer detectable. However, at this time of the digestive period, the phylum *Bacteroidota* dominated 85% of the microbiome of female adults, on average. The observed modulation results in the flavobacterium *E. anophelis* becoming the prevalent component of the microbiome (146,605 reads in blood-fed adults at 48 h of a total of 340,158 reads across all groups; Fig. 1B). This shift is similar to what is observed in the sugar-fed adults at 48 h (44,348 reads assigned to *E. anophelis*), indicating that their proliferation capability may be triggered by the digestion in the midgut environment (Fig. 1). At 48 h, blood-fed adults present a microbial composition like that of sugar-fed groups, except for ascomycete fungi such as *A. flavus* and *M. majus*, restricted to the latter (with 12,508 and 10,260 reads attributed, respectively; Fig. 1B).

Diversity of the microbiomes under different diets

In sugar-fed adults, the total number of microbial phylotypes observed (67 phylotypes) is relatively low throughout the digestive period. Blood-fed adults, conversely, display a high variation of observed phylotypes (average of 136 ± 18 phylotypes) during digestion, reaching the peak of 174 phylotypes 24 h after the blood

meal, as shown in Fig. 2. The Simpson reciprocal index, however, indicates that the peak of microbial diversity in the sugar-fed group is reached at 12 h, decaying in the next 36 h (Fig. 2A). The blood-fed group shows an inverse pattern, with its lowest microbial diversity at 12 h after feeding and increasing diversity up to 48 h (Fig. 2A). The discrepancies between the indices indicate that a higher evenness (i.e. the lack of dominance of one or a few taxa) drives the increase in diversity in sugar-fed adults. The contrary is also true for the blood-fed group, which has less evenness with a few highly abundant species (Fig. 1B). Considering that blood digestion in the midgut of *A. aegypti* mosquitoes lasts 30 to 40 h (31, 32), the microbial overlap in samples from blood- and sugar-fed groups at 48 h indicates that the microbiome of both groups returns to a similar state after digestion. The major modulation of the microbiome is therefore triggered within 24 h after feeding, depending on the type of diet. Further compositional analyses in this work take these digestive periods into account, and henceforth refer to 48 h groups as the "post-digestive" period.

Notably, sample ordinations show divergent beta-diversity patterns (Fig. 2B). The microbiomes of sugar-fed mosquitoes at 12 h and blood-fed at 12 and 24 h are highly unique, forming separate clusters with distinct centroids (Fig. 2B). Nevertheless, there is no clear distinction between the sugar-fed groups at 24 and 48 h and the blood-fed group at 48 h, which can be observed in the biplot as the superposition of their ellipsoids (Fig. 2B, left panel). The scattered distribution of blood-fed individuals analyzed 48 h after feeding caused its centroid to overlap with all individuals in the sugar-fed groups at 48 h, as opposed to other groups of blood-fed individuals (Fig. 2B), demonstrating a convergent shift back to similar compositions after 48 h since their last meal, regardless of the diet. However, when individuals from groups of blood-fed and sugar-fed adults are analyzed independently of the time elapsed after feeding, the two groups are clearly separated (Fig. 2B, right panel), thereby indicating that the type of diet is the best explanatory variable to the compositional microbiome differences ($P < 0.001$; $R^2 = 0.27$), followed by hours post-feeding ($P < 0.001$; $R^2 = 0.16$). These results corroborate previous studies demonstrating that meal sources may directly affect the mosquito microbial community (14, 24, 33).

In the blood-fed group, several phylotypes were shared by the microbial communities of mosquitoes at 12 and 24 h after feeding (117 phylotypes; Fig. 2C). The shared diversity further decreases to 66 phylotypes 48 h post-feeding. However, a peak of unique microbial phylotypes was observed 24 h after the blood meal. These findings confirm the compositional narrowing of the microbial community immediately after the blood meal. In contrast, very few phylotypes were shared among the sugar-fed groups (a maximum of four; Fig. 2C), while unique phylotypes were observed after 12 and 24 h of the sugar meal (25 and five phylotypes, respectively), further elucidating why sugar-feeding is associated with a more diverse microbial repertoire during the earlier stages of digestion.

Microbial network interaction throughout digestion

Microbiomes are complex ecological communities that form interactive networks. To infer these interactions in the microbial community structure, we utilized graphs of microbial richness at the species level, by detecting taxa with significant co-occurrences (Fig. 3). Sugar-fed mosquitoes present fewer microbial co-occurrences than blood-fed adults, but six clusters (S1 to S6) were significantly correlated ($P < 0.05$; $R > 0.75$; Fig. 3A). In blood-fed mosquitoes, six clusters (B1 to B6) were also observed, and the densest network consists of three clusters, where B3 and B4 are composed exclusively of specific *Enterobacterales* phylotypes found in blood-fed adults at 12 and 24 h after feeding, respectively (Fig. 3B). Intriguingly, a third cluster, B5, is composed of microbial taxa that are distantly related, unlike what was observed in the other clusters (Fig. 3B). These results indicate that the structure of *A. aegypti*'s microbial community may not be driven only by the type of diet and digestion periods but also by taxon-specific interactions. The detection of several different species within *Pseudomonas* spp. and *Elizabethkingia* spp. in clusters S4, S5, and B1 is likely a methodological artifact due to the sequence homology of closely related phylotypes within each of these distinct genera. Clusters marked with special characters (*, ○, and #) are highly similar in composition between sugar and blood-fed groups. These microbial phylotypes are detected only in the blood post-digestive period but are evenly distributed among individuals fed with sugar. Although the microbial networks presented in this study are influenced by the strong differences in the microbiome of sugar- vs. blood-fed mosquitoes, Hegde et al (34) showed microbiome networks with a similar composition of bacteria in field-collected mosquitoes, containing *Pseudomonas* spp., as well as unclassified *Enterobacterales* and *Flavobacteriales*. They also reported a prominent co-exclusion of *Enterobacterales* and *Pseudomonales*, confirming that this phenomenon may also occur in wild mosquitoes.

We used part of the microbial richness consistently observed in co-occurrence networks to predict the diet and the time elapsed after a meal. This relationship can be observed in Fig. 4A, which shows that *P. fluorecens*, *A. flavus*, and *M. majus* have the strongest association ($P < 0.05$) with sugar-fed adults at 12 h. On the other hand, the microbiome of the blood-fed adults was dominated by the bacterial phylotype *Raoultella ornithinolytica* (169,307 reads) at 12 h post-feeding, while *Kluyvera intermedia* (196,549 reads) is predominant at 24 h post-feeding. The midgut of individuals metabolizing blood (12 and 24 h) shares a significant association with the phylotypes *Salmonella enterica*, *Enterobacter cloacae*, and other *Enterobacterales* phylotypes. After 48 h, the relative abundance of *E. anophelis* is assigned to both groups of adult mosquitoes regardless of the diet, making this phylotype the most descriptive of the post-digestive state (Fig. 4A). Further analysis of the abundance of *Flavobacteriales* (represented by mainly by *E. anophelis*) and *Enterobacterales* revealed that both orders reach their peak richness opposite to all other taxonomic orders. The order *Flavobacteriales* showed less distinct peaks in the diversity distributions, whereas the *Enterobacterales* displayed only two well-defined richness peaks when no other microorganisms were detected. This finding indicates that the proliferation of these bacteria in the midgut of mosquitoes is mutually exclusive with the presence of other microbial components, given the microenvironmental pressure presented by blood digestion (Fig. 4B). We conducted functional enrichment analyses to delve into the nature of the association of these bacterial lineages with the mosquito.

Functional profiling of the microbiomes during digestion

The functional profile of metagenomes was characterized using the SEED database (35) to analyze the impact of the compositional shifts on the microbiome's functional gene pool. Results indicate that, similarly to what was observed in the microbial compositional analyses, the functional profiles of midguts of blood-fed mosquitoes are strikingly different at 12 and 24 h compared to other adult experimental groups (Fig. 5). Most of the pathways related to the cell wall and ultrastructure (4,395 reads), nucleotide metabolism (44,694 reads), and stress response (1,861 reads in blood-fed

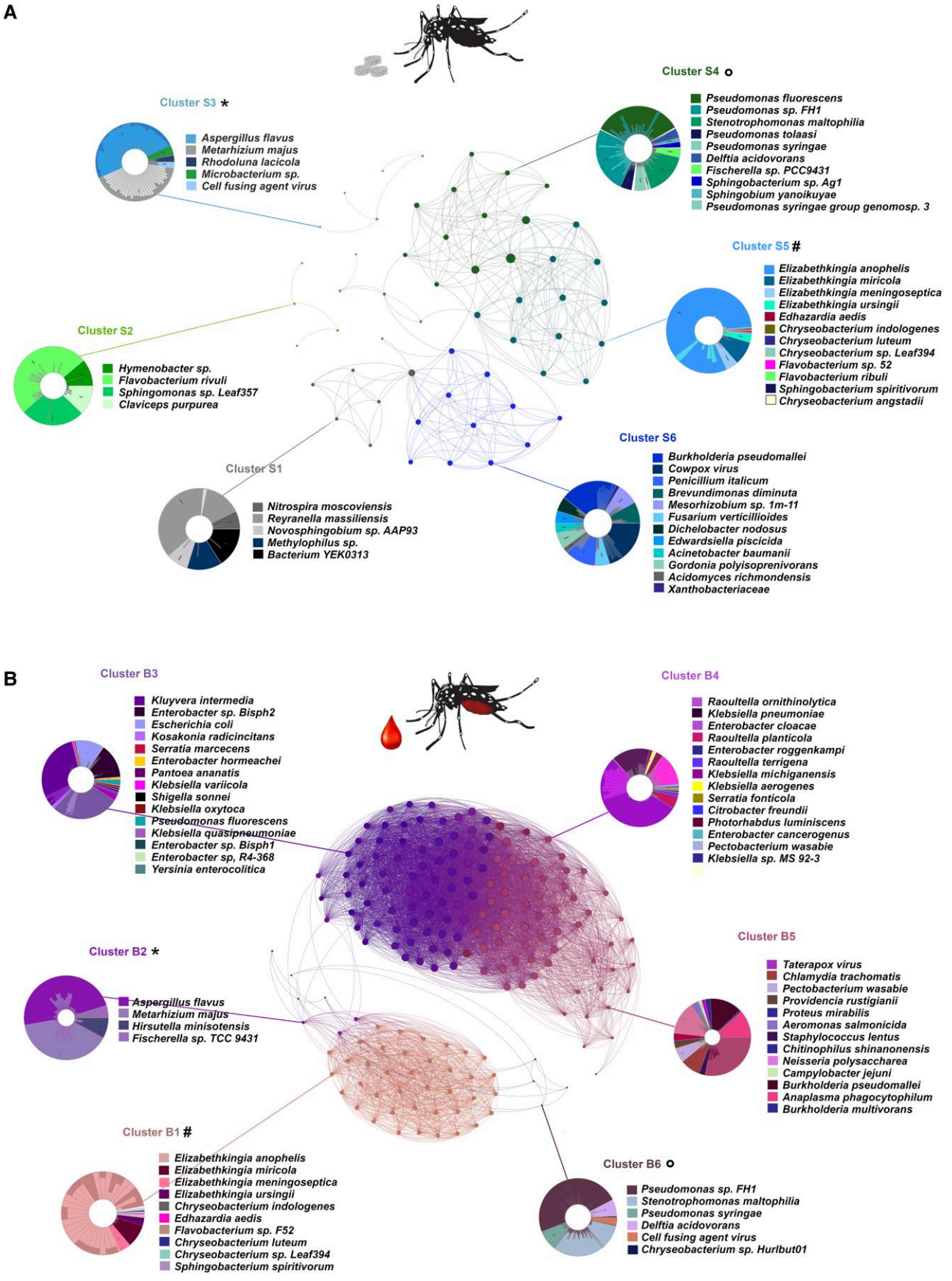

Fig. 3. Co-occurrence networks representing the interactions of the microbial community in adult mosquitoes fed with sugar (A) and blood (B). Graph vertices display microbial species, and edges represent their co-occurrences calculated with the Pearson correlation coefficient ($r > 0.75$; $P < 0.05$). Clusters marked with special characters (*, O, and #) are similar in species composition.

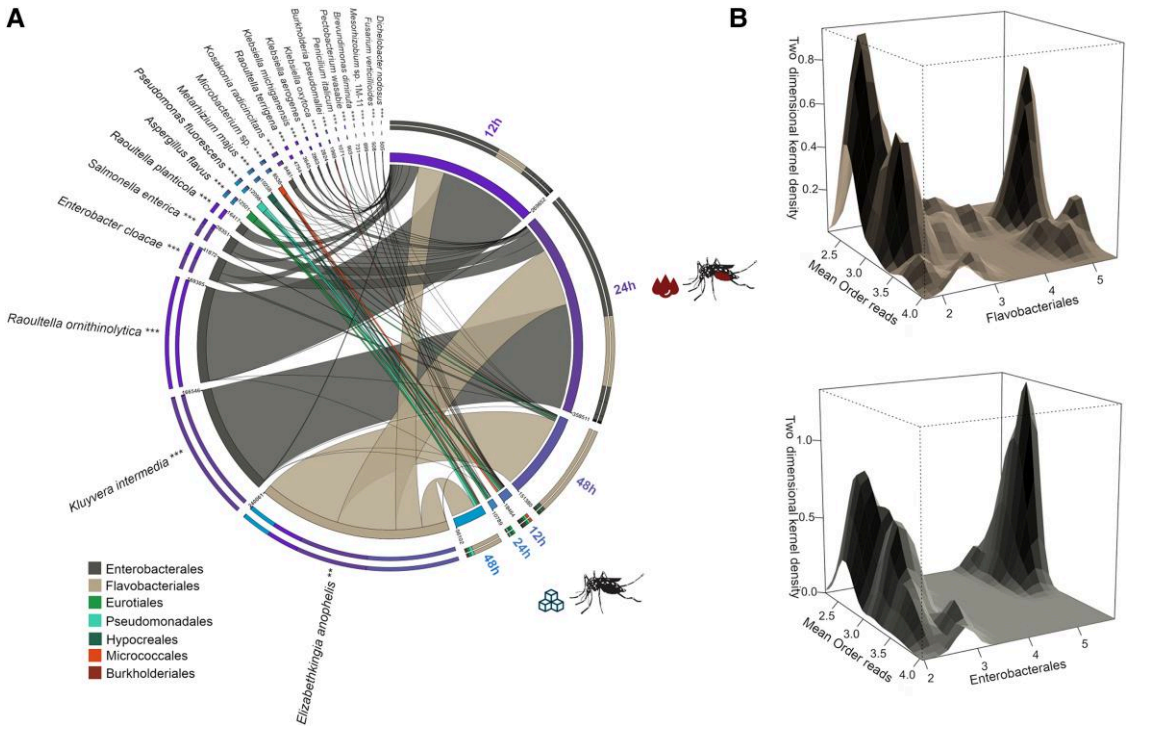

Fig. 4. The landscape of dominant taxa in relation to the microbial composition in the midgut of adult mosquitoes. A) Occurrence of 22 microbial species predictive of digestive states (diet and digestion time). The width of the ribbons indicates the relative abundance in the linear scale. Ribbons link microbial species to experimental groups (represented by the external circles; sugar or blood-fed) and are coded by the taxonomic order. B) Density estimates displaying the occurrence of reads identified in the taxonomic orders *Flavobacteriales* (top) or *Enterobacterales* (bottom) in the x-axis, with the mean of reads attributed to other microbial taxonomic orders in the y-axis, transformed to the log scale.

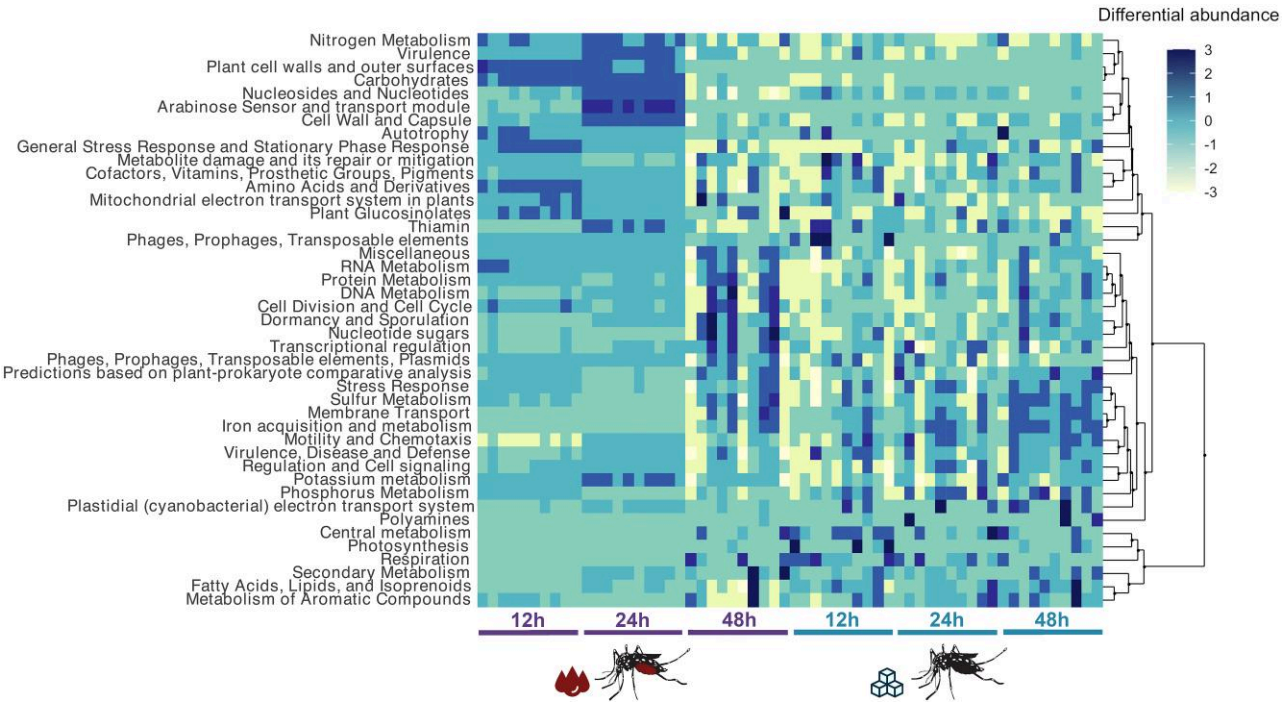

Fig. 5. Functional classification of microbial reads in the midgut of adult females fed with blood and sugar. The heatmap was generated using z-score transformed values acquired in the metabolic pathway enrichment using SEED pathways functional classes in the metagenomes analyzed. Functional pathways are organized by hierarchical clustering. The columns represent individual metagenomes and are grouped according to its experimental group (per-group $n = 10$).

adults). However, 12 h after feeding, fewer genes are related to motility and chemotaxis than in other groups (30,194 reads). After 48 h, the functional profile of blood-fed adults shifts back to a state that resembles those observed in sugar-fed groups (Fig. 5), further showing that the shifts in microbial composition cause changes in the functional pool of the microbiome.

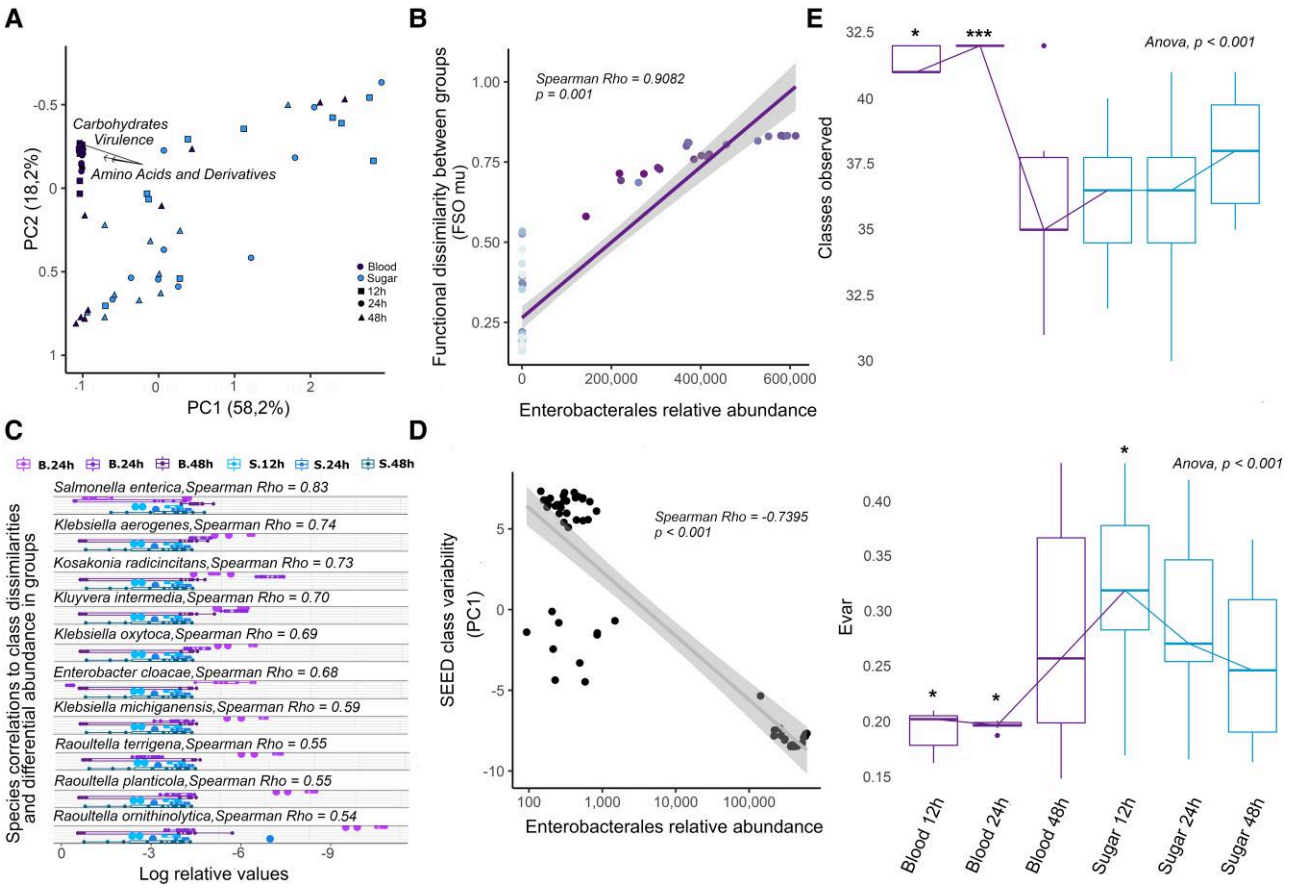

Fig. 6. Functional diversity of the midgut microbiome in adult mosquitoes under different feeding regimes and digestion times. A) Biplot showing a PCoA of Bray–Curtis dissimilarities (PC1 = 58.2% vs. PC2 = 18.2%) displaying SEED pathways responsible for the sample ordination, coded by different diets and shapes, which represent hours post-feeding. B) Fuzzy set ordination (FSO) using a generalized linear model (GLM) to display the correlation between the functional dissimilarity matrix and the relative abundance of *Enterobacterales*. C) Differential occurrence of *Enterobacterales* species classified as predictors for the blood diet in log-relative scale and their individual correlations with the functional dissimilarities. The groups are composed of type of diet and time post-feeding (B.12 h = Blood 12 h post-feeding; B.24 h = Blood 24 h post-feeding; B.48 h = Blood 48 h post-feeding; S.12 h = Sugar 12 h post-feeding; S.24 h = Sugar 24 h post-feeding; S.48 h = Sugar 48 h post-feeding). D) Scatter plot using GLM to display the correlation between the principal explanatory coordinate (PC1) and the relative abundance of *Enterobacterales*. E) Boxplots showing the distribution of observed SEED pathways and their variability indices (Evar). Medians are indicated by the trend line, and the global and pairwise significances are assessed with ANOVA and Wilcoxon's tests, respectively.

Dimensionality reduction demonstrates that reads assigned to carbohydrate metabolism (209,958 reads), amino acids and derivatives (127,903 reads), and virulence (100,710 reads) are responsible for the uniqueness observed in the functional profile of blood-fed mosquitoes at 12 and 24 h (Fig. 6A). Such singularity displays a strong positive correlation with the proliferation of *Enterobacterales* (Fig. 6B), namely *S. enterica* (28,351 reads; $\rho = 0.83$), *Kosakonia radicincitans* (8,480 reads; $\rho = 0.73$), *Kle. aerogenes* (2,863 reads; $\rho = 0.74$) and *Klu. intermedia* (196,549 reads; $\rho = 0.70$). These bacterial phylotypes are the most significantly correlated with the functional enrichment in the blood-fed adults at 12 and 24 h, as shown in Fig. 6C. Additionally, the principal coordinate that best explains the variability of all SEED metabolic pathways (PC1 = 58.2%) has a significant negative correlation with the relative abundance of *Enterobacterales* (Fig. 6D), thereby indicating that these bacteria are likely responsible for narrowing the metabolic spectrum in the mosquito midgut during the digestion of blood, and agree with the functional diversity analysis shown in Fig. 5. This trend can be better observed in Fig. 6E (upper panel), which shows a higher number of metabolic pathways in the midgut of mosquitoes in blood-fed adults at 12 and 24 h than in other groups. Yet, the functional variability (Evar; Fig. 6E, bottom panel) observed in blood-fed adults at 12 and 24 h is lower compared with all other groups, showcasing that, even though many pathways are detected, most of the reads are concentrated in the three enriched pathways. These results indicate that the blood meal triggered the proliferation of specialized opportunistic *Enterobacterales*, who play a more significant role than previously thought in the digestion of blood in the mosquito's midgut. This hypothesis corroborates previous work describing the main association of the mosquito microbiome with gene expression related to metabolic and nutritional pathways (36). Different enterobacteria species have also been associated with an increased digestive capability of fructose in *A. albopictus* (22), and our results describe their potential role in blood digestion in a hematophagous insect. It is possible to speculate that these enterobacteria may be involved in all natural digestive processes of the mosquito in vivo, but not associated with unnatural processes, such as the digestion of sucrose used in our analyses.

To further characterize the modulation of correlated *Enterobacterales* phylotypes in the digestive period (*S. enterica*, *Ko. radicincitans*, *Kle. aerogenes*, and *Klu. intermedia*), the relative abundance of reads assigned to each phylotype, and principal functional pathways are shown for blood-fed adults at 12 and 24 h in Figure S5 ($\rho > 0.70$;

including carbohydrate metabolism, amino acids and derivatives, and virulence). It is possible to observe an overlap of metabolic pathways and microbial abundance of specific *Enterobacterales*, reiterating the strong relationship between the taxonomic and functional classification in the midgut of blood-digesting mosquitoes. Considering the virulence pathway, we detected the presence of functional classes related to multiple antimicrobial resistances, notoriously multi resistance efflux pumps (12,324 reads), and fluoroquinolone resistance (4,743 reads), as well as scavenging pathways (Figure S5C). The detection of commensal microorganisms with antimicrobial resistances has already been reported in the midgut of *A. aegypti* with culture-dependent methods (37). In this study, we document evidence of a putative resistome associated with the microbiome in these mosquitoes.

Under the carbohydrate metabolism pathway, most of the SEED classes are related to the utilization and anabolism of sugars, including maltose and maltodextrin (9,527 reads), and serine utilization in the glyoxylate cycle (15,409 reads). The enrichment of these genes may indicate an increase in the metabolic demand caused by the proliferation of enterobacteria, particularly that of *Klu. intermedia* after blood meals in *A. aegypti*. In fact, previous studies demonstrate that the consumption of dextrose may increase adult mosquitoes' lifespan (38–42). This finding suggests that the blood meal may indirectly influence the hosts' fitness by driving the proliferation of beta-hemolytic enterobacteria, which could alter the availability of macro-nutrients to the host. Lastly, the amino acids and derivatives pathway include functional classes related to glycine and serine metabolism (10,368 reads), as well as biosynthesis (16,643 reads) and degradation (12,018 reads) of methionine (Figure S5C). These pathways have key enzymes for the digestion of blood, which is the case of the glycine and serine utilization pathways detected in the microbiome of adult mosquitoes. The silencing of one of the pivotal enzymes in the latter pathway (serine transferase, SMHT) caused the formation of clots of nondigested blood in female mosquitoes' midguts and a phenotype of ovarian underdevelopment (43), corroborating with our results and strengthening the hypothesis of a nutritional symbiosis between enterobacteria and *A. aegypti*.

Elizabethkingia anophelis genome and phylogenomics

The prevalence of *E. anophelis* throughout and—in high relative abundances—in the post-digestion microbiomes is noteworthy. Initially isolated from *Anopheles gambiae* (44), this gram-negative flavobacterium has been implicated as the causative agent of multiple outbreaks in humans worldwide, raising concerns about the possibility of its vectorial transmission (45–48). Its pathogenesis is characterized by nosocomial bacteremia leading to human sepsis and has been associated with neonatal meningitis, with high mortality rates (49).

Further evidence supporting the dominance of *E. anophelis* in the midgut of *A. aegypti* comes from the recovery of its genome from the reads assigned to *Elizabethkingia* spp. The genome is divided into 210 contigs totaling 4,248,675 bp with an average coverage of 1289-fold (N50 = 35.9 kb; Tables S2 and S3 for detailed quality assessments). The genome size is within the average reported for other *E. anophelis* genomes publicly available (average of 4,101,937 bp), and the annotation identified 4,238 coding sequences and 46 ribosomal RNA subunits. The BUSCO analyses indicated that the MAG presents > 91.4% completeness, with 670 complete genes identified (661 single-copy, nine duplicated), 22 fragmented, and 41 missing when the database for the Flavobacteriales lineage is used for comparison (see Figure S6 for the comparative identification of orthologs using additional lineages databases). A ribosomal multilocus sequence typing (rMLST) resulted in 38 exact matches against 53 bacterial ribosome protein subunits scanned, providing 92% of support for the MAG assignment to *E. anophelis* (Table S4). Further analyses performed with fIDBAC against a curated bacterial type-strain genome database also assigned the MAG as *E. anophelis* based on >98% average nucleotide identity with 95% of the genome aligned (Table S5). Although previous studies using metabarcoding have shown a high abundance of the flavobacterium *Chryseobacterium* sp (34, 50), our results do not indicate the prevalence of this phylotype in our datasets. However, *Chryseobacterium* and *Elizabethkingia* are closely related taxa, and the assignment differences may be explained by the higher taxonomic resolution power of the whole-genome sequencing (WGS) shotgun metagenomics. This confirms that *Elizabethkingia*, in particular *E. anophelis*, is the dominant genus in the midgut microbiome of *A. aegypti*, corroborating with previous findings in mosquitoes from an Australian colony (51).

The comparative phylogenomic analysis corroborated the grouping of the recovered MAG into *E. anophelis* species clade (Fig. 7A). The ML tree recovered *E. anophelis* as monophyletic, and as a sister species to a group formed by (*E. miricola*(*E. bruuniana*(*E. ursingii* + *E. occulta*))) (Fig. 7A). Within *E. anophelis*, the branch lengths are notably short, and the group presents 99% pairwise identity, indicating a recent divergence of the strains. The first diverging strain of *E. anophelis* was the lineage F0708, which was isolated from a human in 1981 according to NCBI records (accession GCA_032842055.1). All remaining 683 *E. anophelis* strains used for phylogenomic reconstruction in this study originate from this lineage and form two main clades, designated here as Clades 1 and 2 (Fig. 7). The MAG recovered was nested into Clade 1, closely related to a group of clinical isolates from humans in the United States, India, Russia, and China (Fig. 7B). Interestingly, the only other strain of *E. anophelis* reported so far in *A. aegypti* (accession GCA_029623655.1) (51); was not grouped with our MAG and is nested within Clade 2 alongside other mosquito-derived strains (Fig. 7B). This clade is a sister group to a set of human-isolated samples. These results suggest that the lineages specifically associated with *A. aegypti* may have emerged at least twice in the evolutionary history of *E. anophelis*. A third event of mosquito colonization by *E. anophelis* was observed in the clade that groups the strains R26 and Ag1, both isolated from the midgut of the malaria mosquito *A. gambiae* (52). From the perspective of the functional role in the host, our results consistently show that both sugar and blood diets interfere with the capacity of proliferation of this bacterium in the midgut. Recent evidence suggests that *E. anophelis* is also found in the saliva and salivary glands of *A. albopictus* (53), corroborating with our phylogenomic results that transmission of human-pathogenic strains by mosquitoes is a possible route. Interestingly, in the same study, colonization by *E. anophelis* was correlated with lower ZIKV titers in mosquito co-infection assays in vivo, which could represent the key to the nature of its symbiosis with *A. aegypti*.

Concluding remarks

The depth of microbiome characterization provided by shotgun metagenomics has unveiled associations between microorganisms and insect hosts, contributing to a better understanding of physiology and metabolism, as well as the roles of the microbiome in vector capacity and the transmission of insect-borne diseases. In this work, we sequenced the microbiomes of 70 individuals of *A. aegypti*, including 10 larvae and 60 adults fed with blood and sugar. The larval microbiome exhibited low diversity and was

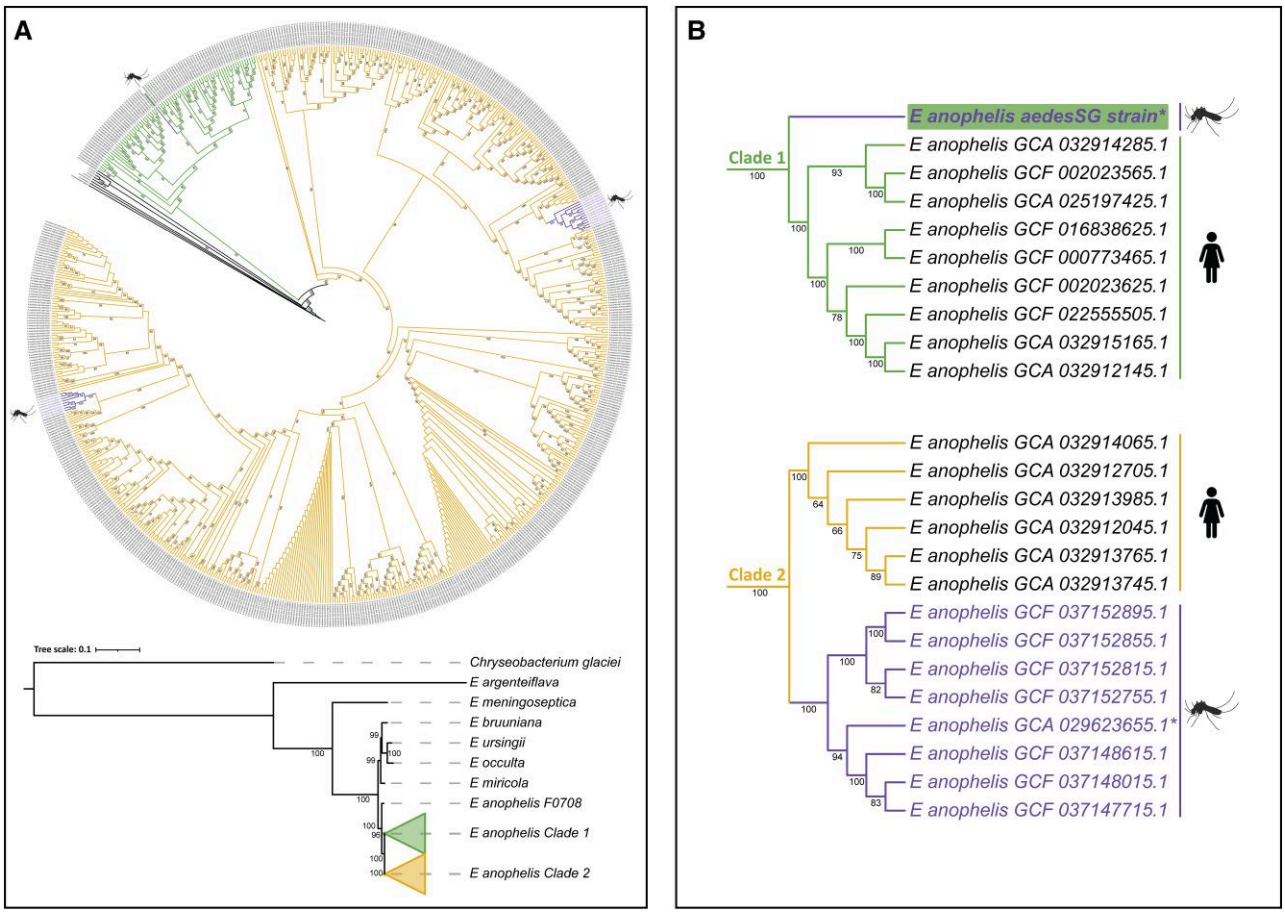

Fig. 7. Phylogenomic analyses of *E. anophelis*. A) ML tree including 690 *Elizabethkingia* genomes and genomes and *Chryseobacterium glacei* as an outgroup, highlighting two main clades within the *E. anophelis* species: Clade 1 and Clade 2. Branches with a mosquito icon denote strains of *E anophelis* associated with mosquito species. The MAG recovered in this study is highlighted. Below the circular tree, a condensed tree shows relationships within the genus *Elizabethkingia*. B) Details of relationships among *E. anophelis* genomes specifically associated with the host *A. aegypti* (genomes are indicated with an asterisk). The genome assembled in this study groups with clinical samples from humans within Clade 1 (tree at the top). The genome isolated from an Australian lineage of *A. aegypti* (tree at the bottom (51)); groups with samples from other mosquitoes within Clade 2 and is recovered as a sister group to samples isolated from humans.

dominated by *Microbacterium* spp. This profile undergoes a notable change during adulthood. Furthermore, after digestion in adults, no significant difference was observed in the microbiome of blood-fed and sugar-fed individuals. In fact, transient changes in microbial composition, diversity, and function were mainly driven by blood meal and lasted only for the digestion period in adults. A significant association of *Enterobacterales* phylotypes with blood digestion was identified, particularly with *Klu. intermedia*. Hematophagy is a habit intimately associated with the mosquito's reproductive cycle, and the proliferation of enterobacteria in response to the blood meal stimulus is likely associated with important metabolic and physiological changes required to deal with oxidative stress and blood digestion. Functional characterization of the microbiome further revealed that potential changes in pathways follow the microbial compositional shifts observed after a blood meal intake. The digestive process also modulates the presence of prominent bacteria, such as *E. anophelis*, which are consistently observed colonizing the midguts of individual mosquitoes during the post-digestive period and may act as potential symbionts of *A. aegypti*. Both *Klu. intermedia* and *E. anophelis* are promising candidates for further assessment of physiological impacts on the mosquito host, potentially serving as control agents for both the vector population and the transmission of arboviruses.

Materials and methods

Mosquito rearing and feeding

Aedes aegypti Singapore strain was reared in a Climatic Test and Plant Growth chamber (Panosnoic MLR-352H) at a temperature of 28 ± 1°C, relative humidity of 80 ± 5%, and a photoperiodic regime of 12:12 h (light:dark). Mosquito breeding followed a previously established protocol (54). In brief, 2–4-week-old *A. aegypti* eggs were hatched in sterile water using a vacuum for 15 min. Newly hatched L1 larvae were bred in a plastic bowl with a density of 2.5 mL/larva in sterile water. Mosquito larvae were fed with a mixture of fish food (TetraMin Tropical Flakers)/brewer's yeast (yeast instant dry blue/Bruggeman) at a ratio of 2:1. The feeding regimen was as follows: 25 mg (day 1), 32 mg (day 2), 56 mg (day 3), 130 mg (day 4), 200 mg (day 5), and 100 mg (day 6). Pupae were collected and placed into a 17.5 × 17.5 × 17.5 cm cage (BugDorm-4S 1515) supplied with one vial of sugar and one vial of sterile water, both of which were replaced twice per week.

Experimental design

For sugar-feeding mosquitoes, 80 newly emerged *A. aegypti* mosquitoes (40 females and 40 males) were maintained in a 17.5 × 17.5 × 17.5 cm cage (BugDorm-4S 1515) supplied with one vial of

sugar and one vial of sterile water. For blood-feeding mosquitoes, 80 newly emerged *A. aegypti* mosquitoes (40 females and 40 males) were maintained in a 17.5 × 17.5 × 17.5 cm cage (BugDorm-4S 1515) supplied with one vial of sugar and one vial of sterile water for three days to allow maturation before being fed with *Sus scrofa domesticus* blood using an artificial membrane feeding system (Hemotek Ltd, UK, or Orinnotech, Singapore). For each group, three subgroups consisting of ten adults were separated based on time elapsed after their last blood (blood-fed adults) or sugar (sugar-fed adults) meal: 12 h (blood-fed $n = 10$; sugar-fed $n = 10$), 24 h (blood-fed $n = 10$; sugar-fed $n = 10$), and 48 h (blood-fed $n = 10$; sugar-fed $n = 10$), with a total of 60 adult individuals. Additionally, ten individuals in the fourth-instar larval stage were analyzed to compare the two developmental stages and assess whether the larval microbiome is transmitted to adults, making up a total of 70 experimental mosquito samples. One sample of the larvae rearing water (W), one sample of the distilled water source used to prepare the water-sugar solution (DW), one sample of the water-sugar solution itself (WS), and one sample of the *Sus scrofa domesticus* blood (B) were analyzed to assess their contribution to the microbiome of mosquitoes. Moreover, we also sequenced three nonmosquito (NM) samples as DNA extraction reagent controls, each corresponding to a different extraction protocol used: (i) NM-1, consisting of a negative control for the kit DNeasy Blood and Tissue (Qiagen) used for mosquito DNA extraction; (ii) NM-2, a negative control for the kit DNeasy Blood and Tissue (Qiagen) used for DNA extraction of minipig blood; and (iii) NM-3, consisting of a negative control for the kit DNeasy PowerWater (Qiagen) used for extracting DNA from water samples. Therefore, a total of 77 samples (Table S1) were sequenced, processed, and analyzed with the same workflow.

DNA extraction and sequencing

Midguts of adult female mosquitoes were dissected in ice-cold PBS using a stereoscope (Olympus SZ61) and kept in 300 µL of phosphate-buffered saline (PBS) 1X, pH 7.4 (Gibco—ThermoFisher). The midguts and larvae were individually macerated with an electric tissue grinder (VWR International), and the homogenates were used for DNA extraction following the insect tissues protocol of the DNeasy Blood and Tissue (Qiagen) kit, according to the manufacturer's instructions. The mini pig blood sample was extracted with the same kit but following the specific protocol for blood DNA extraction. We followed the standard protocol of the DNeasy PowerWater (Qiagen) kit for the water sample extractions. Negative control DNA extractions followed the respective protocols, but no sample was added. DNA yield quantification was performed with the Qubit 1X dsDNA HS Assay Kit (ThermoFisher) in a Qubit 2 fluorometer (ThermoFisher), and DNA integrity was assessed on a Bioanalyzer 2100 system (Agilent) using the High Sensitivity DNA Kit (Agilent). The total DNA for each sample was fragmented using the ultrasonicator Covaris S220 (Covaris Inc.), and the fragments were separated by size in a Pippin Prep electrophoretic system (Sage Science) with 2% agarose gel. Fragments of 300 to 450 bp were collected and purified with Agencourt AMPure XP magnetic beads (Beckman Coulter). Libraries were then built with the Accel-NGS 2S Plus DNA Library Kit (Swift Biosciences), following the manufacturer's protocol. All libraries were indexed with the 2S Dual Indexing Kit (Swift Biosciences), quantified with Quant-iT Picogreen (Invitrogen), and validated by qPCR with the KAPA SYBR FAST qPCR kit (Kapa Biosystems). Equimolar quantities of each indexed library were pooled for multiplex sequencing on the HiSeq 2500 (Illumina Inc.) platform, with a 251 bp paired-end protocol. Sequencing was performed at the Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University (Singapore).

Processing of sequenced datasets

The raw fastq sequencing files were trimmed for both adapter and low-quality sequences using cutadapt v. 1.15 (55). A maximum error rate of 0.2 was allowed to recognize and remove adapters. A quality cutoff of Q20 was used to trim low-quality ends from reads before adapter removal. High-quality reads were aligned against the complete *A. aegypti* genome (GCA_002204515.1) to filter out the host reads. Mapping was carried out using *Bowtie2* (56) with selective parameters for a high sensitivity rate, and reads were filtered with SAMtools (57). The remaining fraction of reads was subsequently translated in six frames and aligned against the NCBI NR protein database release 216 with RapSearch2 v. 2.15 (58) using default parameters. The number of reads generated and analyzed is listed in Table S1.

Taxonomic and functional assignment

After importing alignment results into MEGAN 6 v. 6.18.8 (59), we performed taxa assignment with strict parameters of the Lowest Common Ancestor algorithm, considering the read length generated for each sample with the following settings: Max Expected = 0.01, Top Percentage = 10.0, Min Support = 25, Min Complexity = 0.33, Paired Reads = On. Next, we individually normalized all metagenomes to the dataset with the smallest number of reads of the 70 experimental or seven control samples to obtain the representative relative abundances of assigned microbial taxa. The functional profiles for the different microbiomes were assessed by assigning enriched genes identified to functional classes using the SEED hierarchy (60) database. Results were visualized in a heatmap adjusted to a z-score scale, with a hierarchical grouping of classes.

Diversity estimations and statistical analyses

Species diversity analyses using the Simpson Reciprocal (61) and Shannon-Weaver Indexes (62) were performed in MEGAN 6, while the package vegan v. 2.5–6 (63) was used to generate the chao1 richness index (64) using "species" level for Bacterial, Fungal, and Viral taxa of NCBI taxonomy. Analysis of variance (ANOVA) was performed for each method with 1,000 permutations, and pairwise group significance was assessed with Wilcox's post hoc test. Rarefaction curves were computed with the package iNEXT v. 2.0.20 (65) using parameters of interpolation and extrapolation for Hill numbers (66) to display diversity as a function of the sampling size. Non metrical multidimensional scaling (NMDS) was calculated in vegan, using Bray-Curtis dissimilarity (67) and calculating ellipsoids with confidence intervals based on the centroids for each experimental group. Multivariate analyses with distance matrices (ADONIS) and analysis of similarities (ANOSIM) were employed with 1,000 permutations for the statistical interpretation of the results. The distribution of functional classes was calculated with the Bray–Curtis dissimilarity using a principal coordinates analysis (PCoA). A fuzzy set ordination (FSO) was employed to investigate the functional diversity further using the functional dissimilarities, utilizing the package fso (68). Visualizations were plotted using the package ggplot2 v. 3.3.0 (69).

Distribution of significant species

Linear regressions were directly employed in the dispersion of quantitative variables to assess the correlation of relative species abundance to the observed functional variability in the

metagenomes. To test whether microbial species originate from the same time-diet distribution, we used a phyloseq (70) implementation of the Kruskal–Wallis non parametric test coupled with decision trees of a random forest classifier, in a model adapted from Torondel et al (71). and Ssekagiri et al (72). Briefly, the *P*-values were calculated and corrected for multiple testing using familywise error rate for each pair "phylotype—experimental group." The significance was based on the corrected threshold of $P < 0.05$. Significant species most frequently predicted were then assigned importance in the Random Forest classifier based on the mean decrease in accuracy of their classification (73). To map the probabilistic density of abundant bacterial orders, their kernel bivariate densities (74) were estimated as a function of the median distribution of reads assigned to other microbial orders, and their relationships are visualized in logarithmic scale. Additionally, co-occurrence networks were estimated for the microbiome of adult mosquitoes to detect potential microbial interactions related to different diets and the time of the digestive process. The networks were generated based on an adjacency matrix with a significance threshold $P \leq 0.05$ for each microbial association with a bivariate Pearson correlation coefficient > 0.75 (75), in a model based on Ju et al (76). The attribution of edges was automated with the module iGraph (77), removing vertices with null edges, thereby facilitating the visualization with the software Gephi (78). As the input to these analyses, we generated a new abundance matrix based on the presence of species in at least 10% of the samples, with a minimum of 200 reads, removing likely contaminations and rare taxa to avoid algorithm miscomputations.

Elizabethkingia anophelis genome assembly and *phylogenomics*

All reads assigned to *Elizabethkingia* spp. in the metagenomic analysis were extracted with an in-house script (available at https://github.com/felipe797/Aedes_microbiome) and used as input in SPAdes v. 3.15.2 (79) with the flag "–meta" to perform the assembly of metagenomic sample data in the module metaSPAdes (80). The software MetaBAT v.2 (2, 3) was then used to bin and reconstruct the genome from assembled contigs. The resulting MAG was evaluated using CheckM v. 1.1.6 (81) and QUAST v. 5.2.0 (82) to provide the basic assembly statistics, while BUSCO v. 5.7.0 (83) was used to assess the completeness of the genome with three OrthoDB v10 (84) lineage-specific datasets (Bacteria, Bacteroidetes, and Flavobacteriales). The average coverage was calculated by remapping the extracted reads to the assembly using BBMap v. 38.75 (85). rMLST was performed using the speciesID tool available in the public databases for molecular typing and microbial genome diversity PubMLST (86); (release 2024-04-25) to confirm the metagenomic assignment. The average nucleotide identification (ANI) was performed with fIDBAC (87).

For the phylogenomic analyses, a search for *E. anophelis* genome assemblies was conducted in the NCBI Genomes database, using the filters for "annotated genomes" and "excluding atypical genomes." Genomes with fewer than 1,000 annotated proteins were excluded and whenever there was redundancy in the search (genomes with GCA and GCF accessions), the reference assembly was chosen. The search resulted in 683 *E. anophelis* genomes which were downloaded using the NCBI datasets command-line tools (2024 March 21). Other species of the genus *Elizabethkingia* with available reference genomes were added to the phylogenomic analysis, including *E. argenteiflava* (GCF_009904105.1), *E. bruuniana* (GCF_016599835.1), *E. meningoseptica* (GCF_002022145.1), *E. miricola* (GCF_001483145.1), *E. occulta* (GCF_002023715.1), and *E. ursingii* (GCF_002023405.1). The genome of the species *Chryseobacterium glaciei* (GCA_001648155.1) was used as an outgroup.

The annotation of the *E. anophelis* genome assembled in this study was performed using RASTtk (88), and the annotated proteome was used in the phylogenomic analyses along with the described dataset, totaling 691 proteomes (see Table S6 for the complete list of genomes and accession numbers). The proteomes were used as input for the software Orthofinder v.2.5.5 (89) with the flags "-M msa -oa". Orthofinder was used with the combination of Diamond v.2.1.9.163 (90) for all-versus-all gene comparisons, and MAFFT v. 7 (91) for multiple sequence alignment. A total of 2,545,930 genes analyzed were assigned 8,999 orthogroups, with 604 orthogroups with all species present, and 294 of these consisted of single-copy genes. The concatenated alignment of the orthogroups (385,375 amino acids) was used as input in the program IQTREE v.2.2.0 (92) for the evolutionary model selection with ModelFinder (93) and for maximum-likelihood (ML) tree inference. To reduce computational burden, the best-fit model selection was limited to the models JTT, WAG, and LG and chosen according to the Bayesian Information Criterion. The inference of the ML tree was performed using the evolutionary model JTT + F + R10 with 1,000 replicates of ultrafast bootstrap (94), along with the "-bnni" flag to minimize the impact of severe model violations when using ultrafast bootstrap. The ML tree was analyzed and edited with iTOL v.6.9 (95).

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

This study was partially funded by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro to A.C.M.J. (FAPERJ grant E26/211.473/2021).

Author Contributions

A.C.M.J., Y.C., and S.C.S. designed the study; Y.C., F.G.G., and X.H. performed the breeding and feeding experiments and conducted the midgut dissection. E.L.O. processed the samples and performed DNA extractions. D.I.D.-M. conducted the library preparation and sequencing. B.N.V.P., A.C.M.J., V.K.V., and S.C.S. processed the raw sequencing data. J.F.M.S., B.N.V.P., and A.C.M.J. analyzed the metagenomic data. J.F.M.S. performed the statistical analyses. J.F.M.S. and A.C.M.J. wrote the manuscript with input from B.N.V.P., D.I.D.-M., Y.C., and S.C.S.

Preprints

This manuscript was published on a preprint server: https://doi.org/10.1101/2024.03.04.583003.

Data Availability

Sequencing data and metagenomes of *A. aegypti* will be available at Short Read Archive under the BioProject accession number PRJNA1065965. The specific BioSample and SRA accession numbers for each individual mosquito sample are found in Table S7. All programing scripts associated with this project are available at https://github.com/felipe797/Aedes_microbiome.

References

1 Heiss CN, Olofsson LE. 2019. The role of the gut microbiota in development, function and disorders of the central nervous system and the enteric nervous system. *J Neuroendocrinol*. 31:e12684.
2 Kang DD, *et al.* 2019a. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *Peer J*. 7:e7359.
3 Kang D-W, *et al.* 2019b. Long-term benefit of Microbiota transfer therapy on autism symptoms and gut microbiota. *Sci Rep*. 9:5821.
4 Angleró-Rodríguez YI, *et al.* 2017. An *Aedes aegypti*-associated fungus increases susceptibility to dengue virus by modulating gut trypsin activity. *eLife*. 6:e28844.
5 Cappelli A, *et al.* 2019. Asaia activates immune genes in mosquito eliciting an anti-plasmodium response: implications in malaria control. *Front Genet*. 10:836.
6 Junqueira ACM, *et al.* 2017. The microbiomes of blowflies and houseflies as bacterial transmission reservoirs. *Sci Rep*. 7:16324.
7 Rodríguez-Ruano SM, *et al.* 2018. Microbiomes of North American Triatominae: the grounds for Chagas disease epidemiology. *Front Microbiol*. 9:1167.
8 Vivero RJ, *et al.* 2019. Wild specimens of sand fly phlebotomine *Lutzomyia evansi*, vector of leishmaniasis, show high abundance of *Methylobacterium* and natural carriage of *Wolbachia* and *Cardinium* types in the midgut microbiome. *Sci Rep*. 9:17746.
9 Mason CJ, Raffa KF. 2014. Acquisition and structuring of midgut bacterial communities in gypsy moth (Lepidoptera: Erebidae) Larvae. *Environ Entomol*. 43:595–604.
10 Yun J-H, *et al.* 2014. Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Appl Environ Microbiol*. 80:5254–5264.
11 Douglas EA. 2015. Multiorganismal insects: diversity and function of resident microorganisms. *Ann. Rev. Entomo*. 60:17–34.
12 de Gaio AO, *et al.* 2011. Contribution of midgut bacteria to blood digestion and egg production in *Aedes aegypti* (Diptera: Culicidae) (L.). *Parasit Vectors*. 4:105. https://doi.org/10.1186/1756-3305-4-105
13 Petersen MT, *et al.* 2018. The impact of the age of first blood meal and Zika virus infection on *Aedes aegypti* egg production and longevity. *PLoS One*. 13:e0200766.
14 Almire F, *et al.* 2021. Sugar feeding protects against arboviral infection by enhancing gut immunity in the mosquito vector *Aedes aegypti*. *PLoS One*. 17:9.
15 Apte-Deshpande A, Paingankar M, Gokhale MD, Deobagkaar DN. 2012. *Serratia odorifera* a midgut inhabitant of *Aedes aegypti* mosquito enhances its susceptibility to dengue-2 virus. *PLoS One*. 7:7.
16 Ramirez JL, *et al.* 2014. Chromobacterium csp_P reduces malaria and dengue infection in vector mosquitoes and has entomopathogenic and in vitro anti-pathogen activities. *PLoS Pathog*. 10(10):e1004398.
17 Sharma A, Dhayal D, Singh OP, Adak T, Bhatnagar RK. 2013. Gut microbes influence fitness and malaria transmission potential of Asian malaria vector *Anopheles stephensi*. *Acta Trop*. 128(1):41–47.
18 Aliota M, Peinado S, Velez I, Osorio JE. 2016. The *w*Mel strain of *Wolbachia* reduces transmission of Zika virus by *Aedes aegypti*. *Sci Rep*. 6:28792.
19 Nazni WA, *et al.* 2019. Establishment of *Wolbachia* strain wAlbB in Malaysian populations of *Aedes aegypti* for dengue control. *Curr Biol*. 29(24):4241–4248.e5.
20 Shepard DS, Undurraga EA, Halasa YA, Stanaway JD. 2016. The global economic burden of dengue: a systematic analysis. *Lancet Infect Dis*. 16:935–941.
21 Shragai T, Tesla B, Murdock C, Harrington LC. 2017. Zika and chikungunya: mosquito-borne viruses in a changing world: global change and vectors of chikungunya and Zika. *Ann N. Y. Acad Sci*. 1399:61–77.
22 Scolari F, Casiraghi M, Bonizzoni M. 2019. *Aedes* spp. and their Microbiota: a review. *Front Microbiol*. 10:2036.
23 Muturi EJ, Njoroge TM, Dunlap C, Cáceres CE. 2021. Blood meal source and mixed blood-feeding influence gut bacterial community composition in *Aedes aegypti*. *Parasites Vectors*. 14(1):83. https://doi.org/10.1186/s13071-021-04579-8
24 Muturi EJ, Dunlap C, Ramirez JL, Rooney AP, Kim C-H. 2018. Host blood meal source has a strong impact on gut microbiota of *Aedes aegypti*. *FEMS Microbiol Ecol*. 95(1):fiy213.
25 Foo A, Cerdeira L, Hughes GL, Heinz E. 2023. Recovery of metagenomic data from the *Aedes aegypti* microbiome using a reproducible snakemake pipeline: MINUUR. *Wellcome Open Res*. 8:131.
26 Coon KL, Vogel KJ, Brown MR, Strand MR. 2014. Mosquitoes rely on their gut microbiota for development. *Mol Ecol*. 23:2727–2739.
27 Noskov YA, *et al.* 2021. A neurotoxic insecticide promotes fungal infection in *Aedes aegypti* Larvae by altering the bacterial community. *Microb Ecol*. 81:493–505.
28 Chavshin AR, *et al.* 2015. Malpighian tubules are important determinants of *Pseudomonas* transstadial transmission and longtime persistence in *Anopheles stephensi*. *Parasit Vectors*. 8:36.
29 Chen S, Bagdasarian M, Walker ED. 2015. *Elizabethkingia anophelis*: molecular manipulation and interactions with mosquito hosts. *Appl Environ Microbiol*. 81(6):2233–2243.
30 Salter SJ, *et al.* 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol*. 12:87.
31 Downe AER. 1975. Internal regulation of rate of digestion of blood meals in the mosquito, *Aedes aegypti*. *J. Insect Physiol*. 21:1835–1839.
32 Felix CR, Betschart B, Billingsley PF, Freyvogel TA. 1991. Post-feeding induction of trypsin in the midgut of *Aedes aegypti* L. (Diptera: Culicidae) is separable into two cellular phases. *Insect Biochemistry*. 21(2):197–203.
33 Gonzales KK, *et al.* 2018. The effect of SkitoSnack, an artificial blood meal replacement, on *Aedes aegypti* life history traits and gut microbiota. *Sci Rep*. 8:11023.
34 Hegde S, *et al.* 2018. Microbiome interaction networks and community structure from laboratory-reared and field-collected *Aedes aegypti*, *Aedes albopictus*, and *Culex quinquefasciatus* mosquito vectors. *Front Microbiol*. 9:2160.
35 Overbeek R, *et al.* 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (rsugar-fed adultsT). *Nucleic Acids Res*. 42(Database issue):D206–D214.
36 Hyde J, Correa MA, Hughes GL, Steven B, Brackney DE. 2020. Limited influence of the microbiome on the transcriptional profile of female *Aedes aegypti* mosquitoes. *Sci Rep*. 10:10880.
37 Hyde J, Gorham C, Brackney DE, Steven B. 2019. Antibiotic resistant bacteria and commensal fungi are common and conserved in the mosquito microbiome. *PLoS One*. 14:e0218907.
38 Alvarado WA, Agudelo SO, Velez ID, Vivero RJ. 2021. Description of the ovarian microbiota of *Aedes aegypti* (L) Rockefeller strain. *Acta Trop*. 214:105765.
39 Carter JE, Evans TN. 2005. Clinically significant Kluyvera infections: a report of seven cases. *Am J Clin Pathol*. 123:334–338.
40 de Campos FPF, Guimarães TB, Lovisolo SM. 2016. Fatal pancreatic pseudocyst co-infected by *Raoultella planticola*: an emerging pathogen. *Autopsy Case Rep*. 6:27–31.
41 Posidonio APV, Oliveira LHG, Rique HL, Nunes FC. 2021. The longevity of *Aedes aegypti* mosquitoes is determined by carbohydrate intake. *Arq. Bras. Med. Veterinária E Zootec*. 73:162–168.

42 Singh BR, Singh VP, Agarwal M, Sharma G, Chandra M. 2004. Haemolysins of *Salmonella*, their role in pathogenesis and subtyping of *Salmonella serovars*. *Indian J. Exp. Biol.* 42:303–313.
43 Li X, *et al.* 2019. Serine hydroxymethyltransferase controls blood-meal digestion in the midgut of *Aedes aegypti* mosquitoes. *Parasit Vectors*. 12:460.
44 Kämpfer P, *et al.* 2011. *Elizabethkingia anophelis* sp. Nov., isolated from the midgut of the mosquito *Anopheles gambiae*. *Int J Syst Evol Microbiol*. 61(11):2670–2675.
45 Chew KL, Cheng B, Lin RTP, Teo JWP. 2018. *Elizabethkingia anophelis* is the dominant *Elizabethkingia* Species found in blood cultures in Singapore. *J Clin Microbiol*. 56(3):e01445–e01417.
46 McTaggart LR, *et al.* 2019. Application of whole genome sequencing to query a potential outbreak of *Elizabethkingia anophelis* in Ontario, Canada. *Access Microbiol*. 1:e000017.
47 Perrin A, *et al.* 2017. Evolutionary dynamics and genomic features of the *Elizabethkingia anophelis* 2015 to 2016 Wisconsin outbreak strain. *Nat Commun*. 8:15483.
48 Reed TAN, *et al.* 2020. *Elizabethkingia anophelis* infection in infants, Cambodia, 2012–2018. *Emerg. Infect. Dis*. 26:320–322.
49 Lau SKP, *et al.* 2016. *Elizabethkingia anophelis* bacteremia is associated with clinically significant infections and high mortality. *Sci Rep*. 6:26045.
50 Kim KK, Kim MK, Lim JH, Park HY, Lee S-T. 2005. Transfer of *Chryseobacterium meningosepticum* and *Chryseobacterium miricola* to *Elizabethkingia* gen. nov. As *Elizabethkingia meningoseptica* comb. nov. and *Elizabethkingia miricola* comb. nov. *Int J Syst Evol Microbiol*. 55:1287–1293.
51 Filipović I, Rašić G. 2023. De novo circular genome assembly of *Elizabethkingia anophelis* found in the mosquito *Aedes aegypti* from an Australian colony. *Microbiol Res Announc*. 12(9):e00310–e00323.
52 Kukutla P, *et al.* 2013. Draft genome sequences of *Elizabethkingia anophelis* strains R26T and ag1 from the midgut of the malaria mosquito *Anopheles gambiae*. *Genome Announc*. 1(6):e01030–e01013.
53 Onyango MG, *et al.* 2021. Zika virus and temperature modulate *Elizabethkingia anophelis* in *Aedes albopictus*. *Parasites Vectors*. 14:573.
54 Zhang H, Goh FG, Ng LC, Chen CH, Cai Y. 2023. *Aedes aegypti* exhibits a distinctive mode of late ovarian development. *BMC Biol*. 21:11.
55 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 17:10–12.
56 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 9(4):357–359. https://doi.org/10.1038/nmeth.1923.
57 Li H, *et al.* 2009. The sequence alignment/map format and SAMtools. *Bioinformatics*. 25:2078–2079.
58 Zhao Y, Tang H, Ye Y. 2012. RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinforma Oxf Engl*. 28:125–126.
59 Huson DH, *et al.* 2016. MEGAN community edition—interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Comput Biol*. 12:e1004957.
60 Mitra S, *et al.* 2011. Functional analysis of metagenomes and metatranscriptomes using SEED and KEGG. *BMC Bioinformatics*. 12:S21.
61 Simpson EH. 1949. Measurement of diversity. *Nature*. 163:688–688.
62 Shannon CE. 1948. A mathematical theory of communication. *Bell Syst Tech J*. 27:379–423.
63 Oksanen J, Blanchet G, Friendly M, Kindt R, Legendre P, *et al.* 2019. Vegan: community ecology package. R package Version 2.5-6. https://CRAN.R-project.org/package=vegan.
64 Chao A, Lee S-M. 1990. Estimating the number of unseen species with frequency counts. *Chin. J. Math*. 18:335–351.
65 Hsieh TC, Ma KH, Chao A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods Ecol Evol*. 7:1451–1456.
66 Hill MO. 1973. Diversity and evenness: a unifying notation and its consequences. *Ecology*. 54:427–432.
67 Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr*. 27:325–349.
68 Roberts DW. 2008. Statistical analysis of multidimensional fuzzy set ordinations. *Ecology*. 89(5):1246–1260.
69 Wickham H. 2016. *Ggplot2: elegant graphics for data analysis*. 2nd ed. New York: Use R! Springer International Publishing.
70 McMurdie PJ, Holmes S. 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 8:e61217.
71 Torondel B, *et al.* 2016. Assessment of the influence of intrinsic environmental and geographical factors on the bacterial ecology of pit latrines. *Microb Biotechnol*. 9:209–223.
72 Ssekagiri AT, Sloan W, Ijaz UZ. 2017. Microbiomeseq: an R package for analysis of microbial communities in an environmental context. In ISCB Africa ASBCB Conference, Kumasi, Ghana. https://github.com/umerijaz/microbiomeSeq on 11/2020.
73 Breiman L. 2001. Random forests. *Mach Learn*. 45:5–32.
74 Venables WN, Ripley BD. (2002). *Modern applied statistics with S*, Fourth edition. Springer, New York. https://www.stats.ox.ac.uk/pub/Msugar-fed adultsS4/.
75 Pearson K. 1895. Note on regression and inheritance in the case of two parents. *Proc R Soc Lond Ser. I*. 58:240–242.
76 Ju F, Xia Y, Guo F, Wang Z, Zhang T. 2014. Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. *Environ Microbiol*. 16(8):2421–2432.
77 Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *Inter J Complex Syst* 1695. https://igraph.org.
78 Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring and manipulating networks. In *SanThird International AAAI Conference on Weblogs and Social Media*, San Jose.
79 Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. 2020. Using SPAdes de novo assembler. *Curr Protocol Bioinformat*. 70:e102.
80 Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. 2017. metaSPAdes: a new versatile metagenomic assembler. *Genome Res*. 27(5):824–834.
81 Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. Checkm: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 25(7):1043–1055.
82 Mikheenko A, Prjibelski A, Saveliev V, Antipov D, Gurevich A. 2018. Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*. 34(13):i142–i150.
83 Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol*. 38(10):4647–4654.
84 Kriventseva EV, *et al.* 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucl Acids Res*. 47(D1):D807–D811.
85 Bushnell B. 2014. *BBMap: A Fast, Accurate, Splice-Aware Aligner*. https://escholarship.org/uc/item/1h3515gn on 04/2024.
86 Jolley KA, *et al.* 2012. Ribosomal multilocus sequence typing: universal characterization of bacteria from domain to strain. *Microbiology*. 4:158.
87 Liang Q, *et al.* 2021. fIDBAC: a platform for fast bacterial genome identification and typing. *Front Microbiol*. 12:723577.

88 Brettin T, *et al.* 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep.* 5(1):8365.
89 Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 20:238.
90 Buchfink B, Reuter K, Drost H-G. 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nature Methods*. 18(4):366–368. https://doi.org/10.1038/s41592-021-01101-x.
91 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*. 30(4):772–780.
92 Minh BQ, *et al.* 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 37(5):1530–1534.
93 Kalyaanamoorthy S, Minh B, Wong T, von Haeseler A, Jermiin LS. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods*. 14:587–589.
94 Minh BQ, Nguyen MAT, von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. *Mol Biol Evol*. 30(5):1188–1195.
95 Letunic I, Bork P. 2024. Interactive tree of life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. *Nucl Acids Res*. 52(W1):W78–W82.